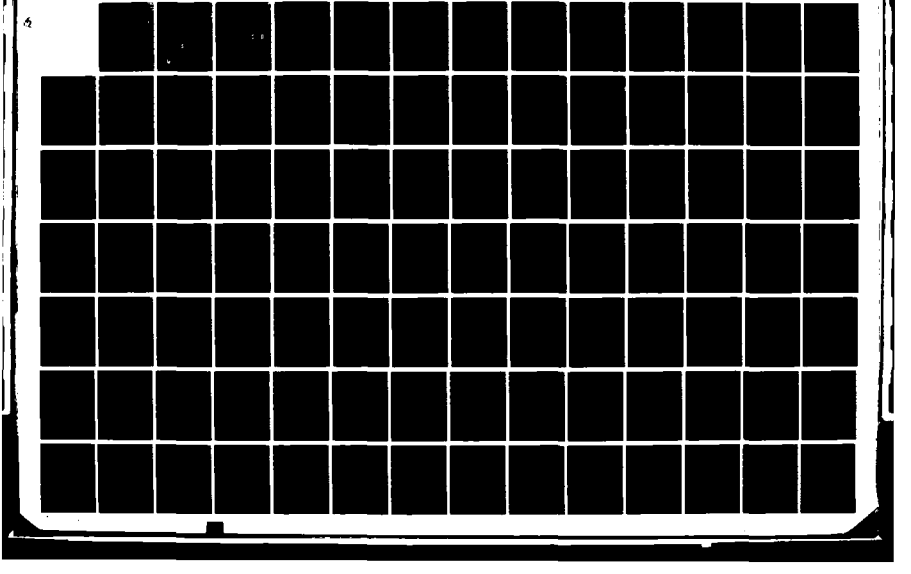
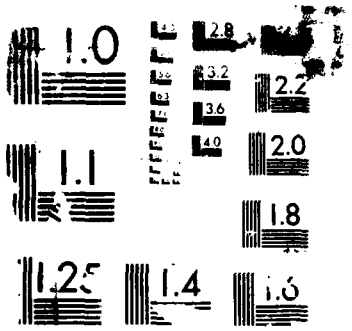
AD-A189 130 ESTIMATION-CORRELATION MODELING AND IDENTIFICATION IN ADAPTIVE ARRAY PROC..(U) PENNSYLVANIA STATE UNIV STATE COLLEGE APPLIED RESEARCH LAB J A TAGUE ET AL. OCT 87 1/2

UNCLASSIFIED ARL/PSU/TR-87-005 N00024-85-C-6041 F/G 17/11 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.8
3.2
3.6
4.0

Applied Research Laboratory

(4)

Technical Report

ESTIMATION-CORRELATION, MODELING, AND IDENTIFICATION IN ADAPTIVE ARRAY PROCESSORS

by

J. A. Tague & L. H. Sibul

PENNSTATE

1855

The Pennsylvania State University
APPLIED RESEARCH LABORATORY
P.O. Box 30
State College, PA 16804

ESTIMATION-CORRELATION, MODELING, AND IDENTIFICATION IN ADAPTIVE ARRAY PROCESSORS

by

J. A. Tague & L. H. Sibul

Technical Report No. 87-005
October 1987

Supported by:
Naval Sea Systems Command

L. R. Hettche, Director
Applied Research Laboratory

Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT: (A) Unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): TR-87-005	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: Applied Research Laboratory, The Penna. State University; 6b. OFFICE SYMBOL (If applicable): ARL	7a. NAME OF MONITORING ORGANIZATION: Naval Sea Systems Command, Department of the Navy
6c. ADDRESS (City, State, and ZIP Code): P. O. Box 30, State College, PA 16804	7b. ADDRESS (City, State, and ZIP Code): Washington, DC 20362
8a. NAME OF FUNDING / SPONSORING ORGANIZATION: Naval Sea Systems Command; 8b. OFFICE SYMBOL (If applicable): NAVSEA	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N-00024-85-C-6041
8c. ADDRESS (City, State, and ZIP Code): Department of the Navy, Washington, DC 20362	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
ESTIMATION-CORRELATION, MODELING, AND IDENTIFICATION IN ADAPTIVE ARRAY PROCESSORS (Unclassified)

12. PERSONAL AUTHOR(S)
J. A. Tague and L. H. Sibul

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Ph.D. Thesis	FROM _____ TO _____	October 1987	168

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Adaptive signal processing, estimator-correlator processing, Karhunen-Loeve expansion, kernel estimation, signal detection, signal estimation

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This dissertation studies several issues involved with implementing adaptive array processors for signal detection and estimation. The basic structure which is examined is the estimator-correlator processor. It establishes a fundamental connection between detection and estimation theory, and solves the detection problem for the generalized exponential class of signal and noise distributions.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT: ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code): 814/865-6344; 22c. OFFICE SYMBOL

System modeling and identification represent two of the central issues studied in this dissertation. Stochastic operator theory and results from functional analysis allow these problems to be solved for a wide variety of random scattering media. The models are based on matrix representations of bounded, linear operators.

The integral kernels needed to implement the estimator-correlator are expressed in terms of Karhunen-Loève expansions. We assert they provide the theoretical means to solve the processor equations, give considerable insight into their mathematical structure, and establish a link between theory and practical realization. Calculating the conditional mean of the channel output requires modeling and identifying the stochastic scattering channel. A Karhunen-Loève expansion of its random Green's function representation is a fundamental approach to stochastic system modeling and identification.

The numerical issues involved with implementation are studied in detail. We show how the processor equations can be solved using robust, state-of-the-art numerical algorithms. The estimator branch is implemented with the CS decomposition. ~~Several interesting relationships~~

ABSTRACT

This dissertation studies several issues involved with implementing adaptive array processors for signal detection and estimation. The basic structure which is examined is the estimator-correlator processor. It establishes a fundamental connection between detection and estimation theory, and solves the detection problem for the generalized exponential class of signal and noise distributions.

System modeling and identification represent two of the central issues studied in this dissertation. Stochastic operator theory and results from functional analysis allow these problems to be solved for a wide variety of random scattering media. The models are based on matrix representations of bounded, linear operators.

The integral kernels needed to implement the estimator-correlator are expressed in terms of Karhunen-Loève expansions. We assert they provide the theoretical means to solve the processor equations, give considerable insight into their mathematical structure, and establish a link between theory and practical realization. Calculating the conditional mean of the channel output requires modeling and identifying the stochastic scattering channel. A Karhunen-Loève expansion of its random Green's function representation is a fundamental approach to stochastic system modeling and identification.

The numerical issues involved with implementation are studied in detail. We show how the processor equations can be solved using robust, state-of-the-art numerical algorithms. The estimator branch is implemented with the CS decomposition. Several interesting relationships

between the Karhunen-Loève basis and canonical matrix decompositions are established.

In summary, ideas from system modeling, identification, detection and estimation theory, and numerical analysis are combined in order to implement optimal array processors for signal detection and estimation.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF FIGURES

ACKNOWLEDGEMENTS

The author wishes to thank Dr. Leon H. Sibul for his guidance and suggestions throughout the work.

The author would also like to thank the committee members for their reading of this manuscript and their constructive comments.

The author would like to thank Dr. Charles Van Loan of Cornell University for copies of his generalized singular value decomposition algorithms.

The author is grateful to Mrs. Phyllis McGarvey for her heroic efforts during the typing of the manuscript.

This work was performed at the Applied Research Laboratory of The Pennsylvania State University, and supported by the Office of Naval Research Fundamental Research Initiatives Program. Their support is most gratefully acknowledged.

Chapter 1

INTRODUCTION

1.1 Problem Statement and Its Importance

1.1.1 General Problem Statement

This dissertation studies several issues arising in implementing adaptive array processors for signal detection and estimation. We shall focus our attention on detecting signals that have propagated through randomly time-varying scattering media. The stochastic nature of the medium causes the returned signal energy to be a random process spread with respect to range, angle, and Doppler. These effects will be modeled by linear, bounded stochastic operators acting on the transmitted signal. An operator theoretic approach to the detection problem provides the means to solve for the optimal processor for a wide class of propagation and scattering channels.

One can show that the optimal array processor structure can be implemented in an "estimator-correlator" structure (Figure 1-1) [1] [2]. In other words, the array measurements are directed into two branches, each of which is described by a matrix filter acting on the data. The matrix filters $\underline{Q}(\cdot,\cdot)$ and $\underline{G}(\cdot,\cdot)$ are found by solving the following integral equations:

$$\int_T \underline{R}_N(t, u)\, \underline{Q}(u, z)\, du = \delta(t - z)\, \underline{I} \qquad (1.1\text{-}1)$$

$$\int_T \underline{R}_1(t, u)\, \underline{G}(u, z)\, du = \underline{R}_y(t, z) \qquad (1.1\text{-}2)$$

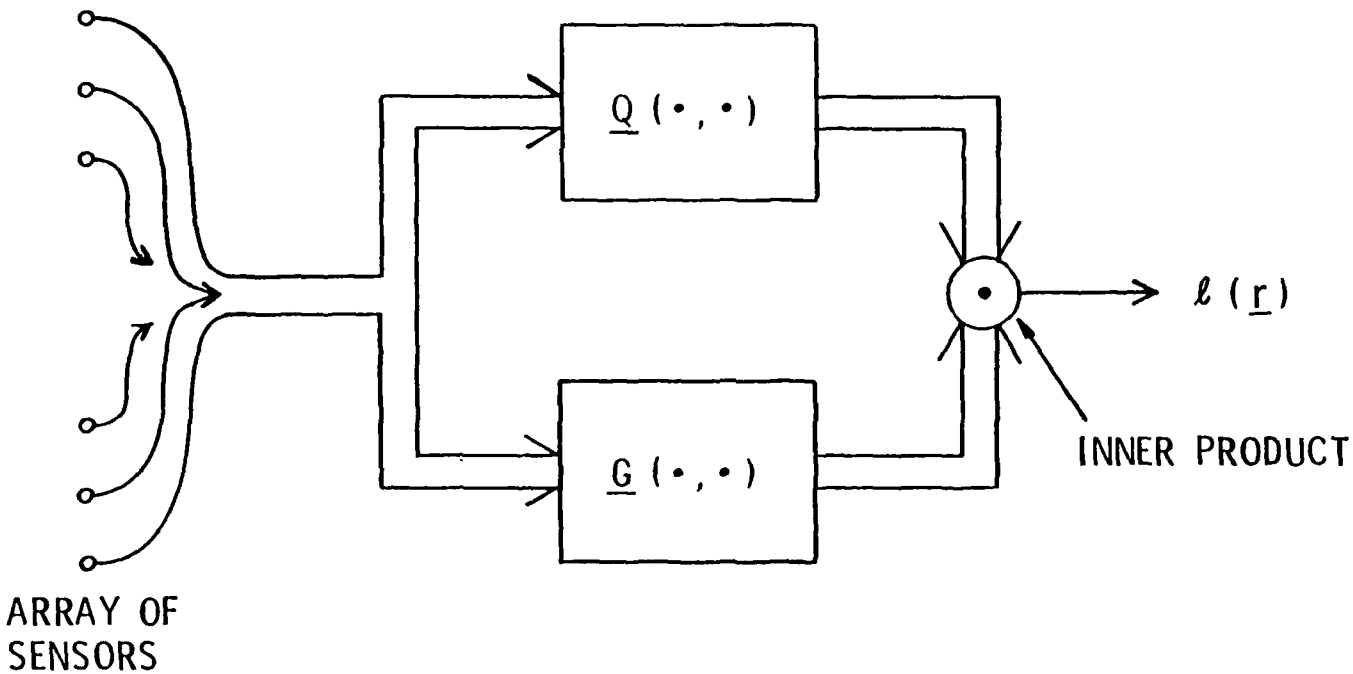


Figure 1-1. The Estimator-Correlator Canonical Structure

where $\underline{R}_N(\cdot,\cdot)$ is the covariance kernel of the measurement noise, $\underline{R}_y(\cdot,\cdot)$ is the covariance kernel of the returned signal alone, $\underline{R}_I(\cdot,\cdot)$ is the covariance kernel of combined returned signal plus noise, T is the observation interval, and $\underline{I}$ is the identity matrix. The lower branch containing $\underline{G}(\cdot,\cdot)$ will be called the estimator branch, because it calculates an optimal estimate of the backscattered signal based on the array data $\underline{r}(\cdot,\cdot)$. The upper branch will be referred to as the inverse filter branch, because it represents the inverse of $\underline{R}_N(\cdot,\cdot)$. The outputs of each branch are correlated to form the scalar-valued likelihood ratio ℓ.

In principle, there is no reason why the optimal detector could not be implemented after solving these equations. However, several problems will become apparent as we consider in detail how to implement this processor in a practical working environment. Solving the equations is so difficult that suboptimal schemes are almost always used.

1.1.2 Specific Problem Statement

The specific issues which shall be studied in this dissertation can be summarized as follows:

(1) To study the relationships among detection theory, estimation theory, stochastic system modeling, and system identification within the context of the estimator-correlator processor.

(2) To establish the connections among Karhunen-Loève expansions, system models, generalized Fourier series, canonical matrix decompositions, and generalized eigenvalue matrix decompositions.

(3) To exploit the preceding results in order to implement the estimator-correlator structure with state-of-the-art computational algorithms.

1.1.3 Importance of the Problem

Most adaptive array processor designs are based on oversimplified descriptions of the backscattering object and transmission medium. They are called point scatterer models, and although they assume random amplitude and phase, their essential features are deterministic. To be more precise, the backscattered signal is represented as a time- and Doppler-shifted version of the original signal which arrives at the array from one specific direction. The return can be interpreted as a point in a three-dimensional space parameterized by time delay, Doppler shift, and arrival angle. Implementing these processors is straightforward.

However, a deterministic treatment of transmission channel effects is rarely an adequate representation of reality. Since many uncontrolled factors determine real wave motion, stochastic descriptions are usually more appropriate. Stochastic media cause the returning signal energy to be distributed, or spread, over random intervals of time delay (range), Doppler, and arrival angle. One can think of the returned energy as distributed over a volume in the previously defined parameter space. The total signal energy available for processing is recovered by integrating over the entire volume.

This discussion leads to the importance of this work. When a processor designed for point channels is employed in a stochastic environment, only a fraction of the returned signal energy is actually processed. As a result, the overall performance of the detector is significantly reduced.

Total signal energy processed determines array processor performance. In order to maximize performance, all available signal energy must be used. This is why optimal structures must be implemented, especially in low signal-to-noise ratio environments. The preceding comments provide the underlying motivation for the work presented in this dissertation.

1.2 Problem Formulation

A schematic diagram of the working environment is shown in Figure 1-2. A signal $s(\cdot)$ is transmitted into the medium over a finite time interval. $s(\cdot)$ shall be called the probing signal, and we will assume its characteristics are known and subject to our control. If a reflecting object is present in the medium, then energy from the probing signal scatters off the object in different directions. A collection of sensors called the array is immersed in the medium and is designed to be sensitive to the signal energy reflected from the object.

All signals not related to the backscattered probing signal represent undesired interference, and will be called noise. It has two components: system noise which arises within the processor, and ambient noise which enters the system through the array. The ambient noise is spatially distributed, but not necessarily isotropically. We shall combine the internal and external noises into a single process modeled by a zero-mean, Gaussian distributed random vector.

The total output of the array is measured over a time interval T and a spatial aperture A. Our objective is to design a processing system which takes the array data and decides if backscattered signal

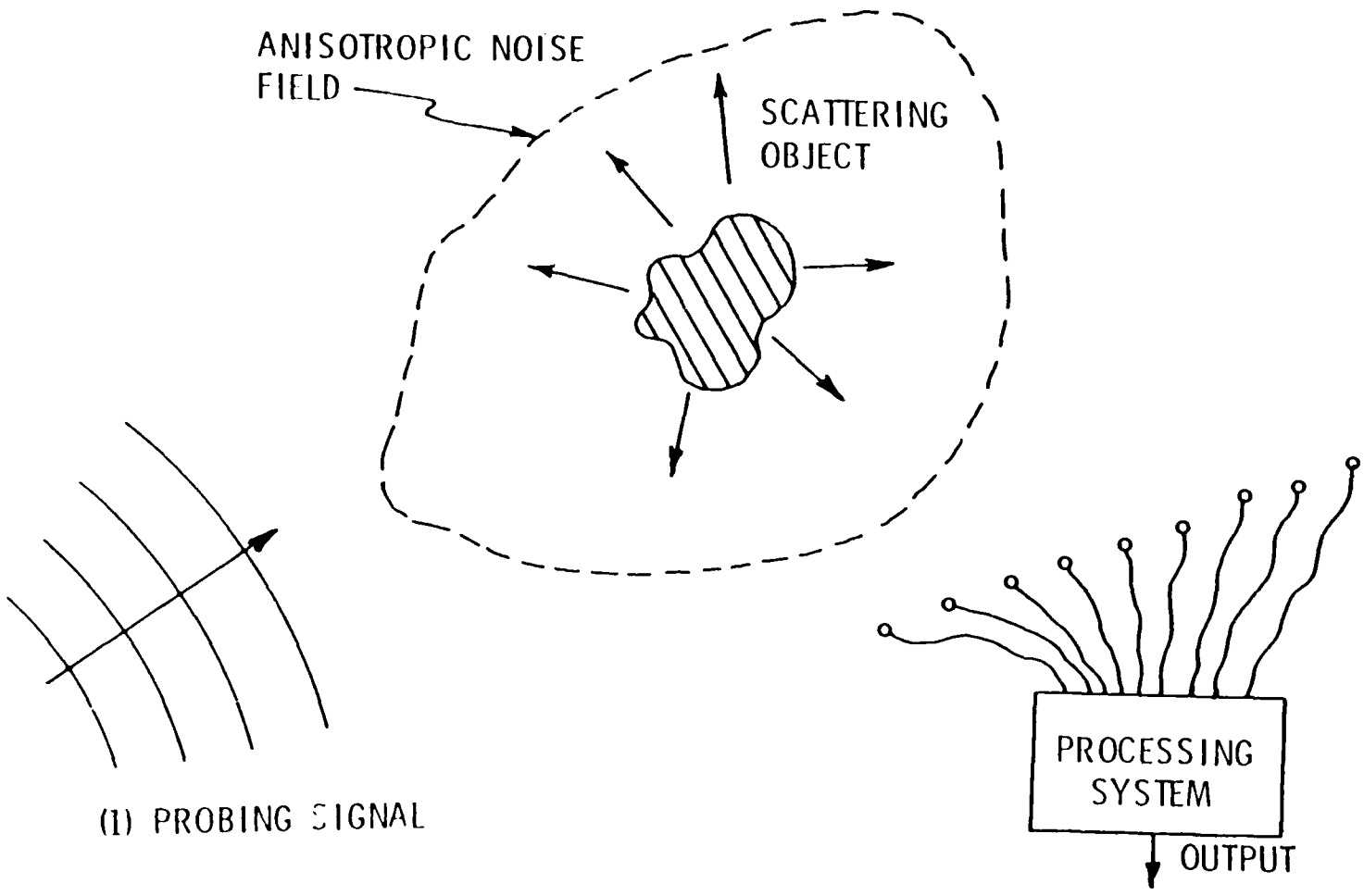


Figure 1-2. System Under Consideration

energy is absent or present in the measurement, in other words, if the object is absent or present.

This problem can be formulated in mathematical terms as follows. The array output shall be denoted by $\underline{r}(\cdot,\cdot)$. It is a complex-valued vector stochastic process, and the i-th element of $\underline{r}(t, \omega)$ represents the measured value of the i-th sensor in the array at time t.

The background and measurement noise entering the processing system will be denoted by $\underline{n}(\cdot,\cdot)$. $\underline{n}(\cdot,\cdot)$ is a zero-mean, Gaussian distributed stochastic process.

The energy backscattered off the reflecting object will be denoted by $\underline{f}\underline{s}(\cdot)$, or by $\underline{y}(\cdot,\cdot)$. This process is inherently random even though $s(\cdot)$ is known; therefore, it must be modeled as a stochastic process as well. $f(\cdot)$ represents the effects of the scattering object and propagation channel. We shall refer to them generically as "the channel," even though many transformations may actually have taken place. $f(\cdot)$ is much more than a convenient shorthand notation. It will be discussed at length later.

Assuming the medium and array are linear, when backscattered energy is present in the measurement:

$$\underline{r}(t, \omega) = \underline{f}\underline{s}(t) + \underline{n}(t, \omega)$$

When the object is absent:

$$\underline{r}(t, \omega) = \underline{n}(t, \omega)$$

In practice, we do not know which form of $\underline{r}(\cdot,\cdot)$ is correct. The process of deciding between these alternatives in a statistically meaningful sense is called signal detection. Implementing the structure which solves the detection problem is the basic issue addressed in this dissertation.

1.3 Basic Approach to the Problem

Let us begin by describing the measurement model in greater detail, since it represents a central part of the overall approach to the problem. We stated that $\underline{f}(\cdot)$ represents the effects of the scattering channel; however, we have yet to define $\underline{f}(\cdot)$ in mathematical terms, or to motivate its usefulness in this context.

$\underline{f}(\cdot)$ is a bounded, linear, stochastic operator which will represent the overall effects of the medium. This approach is not new. It was suggested by Middleton thirty years ago [3] and reintroduced more recently by Sohie [4]. However, we are the first to apply stochastic operator theory systematically to the channel modeling and identification problem.

A stochastic operator theoretic approach is useful for several reasons. First of all, it is clear that deterministic representations of real scattering media are totally inadequate. Stochastic transformations must be used in order to develop realistic models for the channel effects, and stochastic operator theory provides the means to do so. In this dissertation, the stochastic operator modeling problem will be studied in some detail. Also, although an operator theoretic approach to the problem is actually quite abstract, it allows the detection

problem to be solved generically for a wide class of stochastic media, bypassing a detailed mathematical analysis of many special cases. Moreover, this allows theorems and ideas from functional analysis and Hilbert space theory to be used to obtain new results and new insights into the modeling issues.

Now that the nature of $\underline{f}(\cdot)$ has been described in detail, let us turn our attention to solving the detection problem. We call upon the array processing system to make a judgement concerning the nature of $\underline{r}(\cdot,\cdot)$. This problem is distinguished by the facts that the processor has only limited knowledge of the backscattered signal, and that random noise always obscures the signal to a varying degree. Hence, it is logical to conclude that the required judgement must be a statistical inference based on results from statistical decision theory.

The signal detection problem is equivalent to a statistical hypothesis testing problem, in which the hypothesis that noise alone is present is to be tested against the hypothesis that signal and noise are present. These alternatives are expressed in statistical terms by:

$$H_1: \quad \underline{r}(t, \omega) = \underline{f}\underline{s}(t) + \underline{n}(t, \omega)$$

$$H_0: \quad \underline{r}(t, \omega) = \underline{n}(t, \omega)$$

where H_1 and H_0 are abbreviations for hypothesis one and the null hypothesis respectively.

The solution to the detection problem is well known and can be found in any one of several references [1] [3]. The optimal test, or optimal processor, is prescribed by calculating a real-valued statistic

of $\underline{r}(\cdot,\cdot)$ called the likelihood ratio, and comparing this result to a predetermined number γ. If the likelihood ratio exceeds this threshold, we assert that H_1 is true; otherwise, we assert that the null hypothesis is correct.

The processor can be implemented using several equivalent structures. We have already described the "estimator-correlator" realization, in which the non-causal conditional mean estimate of $\underline{s}(\cdot)$ is correlated with a filtered version of the data $\underline{r}(\cdot,\cdot)$ to obtain the likelihood ratio. This is the basic structure that will be studied in this dissertation.

There are several reasons why we have chosen the estimator-correlator structure. First, it establishes a basic connection between detection theory and estimation theory. This can be seen by comparing this structure with the structure solving the known signal in Gaussian noise detection problem (Figure 1-3). Clearly, the two are similar. The estimator-correlator treats the conditional mean estimate as if it were deterministic in the subsequent correlation operation.

While the preceding discussion points out a theoretically elegant connection between the estimator-correlator and other processors, there are more fundamental reasons why an estimator-correlator realization is useful. They can be seen by examining the nature of the solution in greater detail.

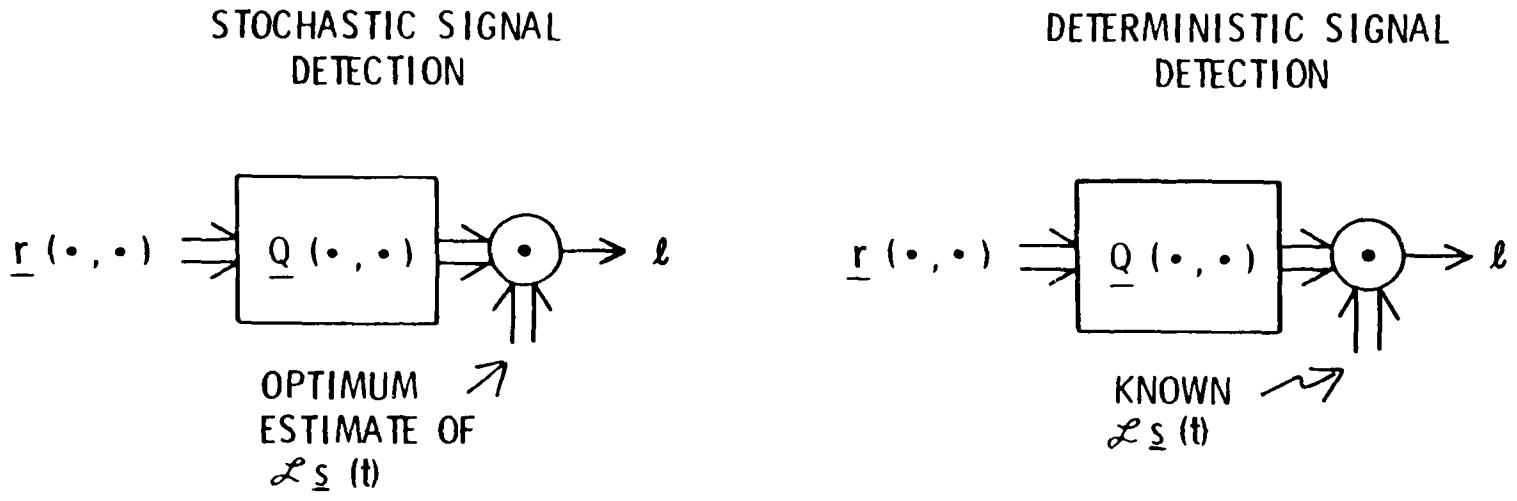


Figure 1-3. Deterministic and Stochastic Signal Detection

Solving Equations (1.1-1) and (1.1-2) for $\underline{G}(\cdot,\cdot)$ and $\underline{Q}(\cdot,\cdot)$ presupposes perfect a priori knowledge of covariance kernels $\underline{R}_N(\cdot,\cdot)$, $\underline{R}_y(\cdot,\cdot)$, and $\underline{R}_I(\cdot,\cdot)$. However, in problems of practical interest, this information is seldom available in advance. This clearly suggests attempting to implement the processor using adaptive signal processing techniques; indeed, adaptation is the key to optimum receiver implementation [5]. The estimator-correlator structure is particularly useful from this point of view.

We will show that $\underline{R}_N(\cdot,\cdot)$ and $\underline{R}_I(\cdot,\cdot)$ can be estimated directly from array data. On the other hand, $\underline{R}_y(\cdot,\cdot)$ can not be obtained directly, since the return $\underline{fs}(\cdot)$ is always obscured by background and measurement noise. If a priori knowledge of $\underline{R}_y(\cdot,\cdot)$ is unknown or incomplete, another means must be found to estimate its salient features in conjunction with the detection process.

This aspect of the problem will be approached from a system modeling and identification point of view. The stochastic transmission media, represented generically by linear stochastic operator $f(\cdot)$, can be interpreted as an unknown linear system. By exploiting results from functional analysis and stochastic operator theory, relationships between $\underline{R}_y(\cdot,\cdot)$ and $f(\cdot)$ suitable for digital signal processing applications can be derived. In particular, the formulations are based on the matrix representation of $f(\cdot)$ and the spectral representation of $\underline{R}_y(\cdot,\cdot)$. Integrating these results into the estimator-correlator structure is straightforward and establishes a fundamental connection

among detection theory, estimation theory, system modeling, and system identification theory.

In this dissertation, we will also study the implementation problem in some detail. The theoretical analysis and practical implementation are both based on Karhunen-Loève expansions of the received data. Of course, the use of series expansions as a theoretical tool in such disciplines as detection theory and estimation theory is well established, and they provide a great deal of insight into the nature of the optimal processor. However, we assert that they are the key to implementation as well. Going beyond formal manipulation of infinite series, we show how these representations lead to structures which can be implemented with state-of-the-art computational algorithms. The results establish relationships among Karhunen-Loève expansions, matrix decompositions such as singular value decompositions and QR factorizations, and generalized eigenvalue problems.

In conclusion, our approach to the problem is based on the estimator-correlator canonical structure. Within this framework, the interrelations among system modeling, stochastic operator theory, identification, and orthonormal expansions will be established. The solutions to the implementation problem are based on Karhunen-Loève series expansions. They lead to structures that can be realized through robust numerical algorithms.

1.4 Review of Previous Work

The estimator-correlator structure was first studied by Price [6] in the context of the Gaussian signal in Gaussian noise detection problem. He noticed that the solution could be interpreted as a modification of the standard "correlation" receiver which solves the known signal in Gaussian noise detection problem. The modification is elegant and intuitive. Specifically, the optimal estimate of the signal process is used as if it were deterministic in the subsequent correlation operation.

Later, Kailath [7] and Esposito [8] studied this structure in more detail. Both argued that it should be optimum or close to optimum for detecting random signals (not necessarily Gaussian) in additive Gaussian noise. However, their results must be interpreted carefully. For example, although Esposito was able to show that an "estimator-correlator" structure exists for a broad class of random signals, it is not possible to interpret the signal estimate as "optimal" in a meaningful sense except for Gaussian signals. Moreover, since the signal process estimates can not be determined uniquely, they actually have little intrinsic value except in the context of the receiver structure itself.

A few years later, Kailath returned to the problem of detecting non-Gaussian stochastic signals $y(\cdot,\cdot)$ in additive Gaussian noise [9], and proved that the log-likelihood functional has the form:

$$\ell(r) = \int_T r(t, \omega)\, \hat{y}(t, \omega)\, dt - \frac{1}{2} \int_T |\hat{y}(t, \omega)|^2\, dt \qquad (1.4\text{-}1)$$

where $\hat{y}(\cdot,\cdot)$ is the causal conditional mean estimate of $y(\cdot,\cdot)$ given $r(\cdot,\cdot)$, assuming that H_1 is true. Equation (1.4-1) shows that an estimator-correlator interpretation of detection can be generalized to encompass a broader class of signals. However, the "correlation" integral has to be defined as an Itô stochastic integral, making practical implementation of (1.4-1) problematical. Furthermore, Kailath pointed out that this result was not merely a question of rigor. Other definitions for the correlation integral yield detector structures inconsistent with previously obtained results.

More recently, Schwartz extended these ideas to discrete time problems [10], and in addition, demonstrated that the structure is optimal when the data come from generalized exponential distributions, a broad and important class of distributions. During this research, the author generalized his results to the vector measurement case. Once again, a conditional mean estimate of the signal is the central part of the structure, and it appears in a correlation integral used to calculate the likelihood ratio. However, evaluating the likelihood ratio is quite difficult, and as Schwartz himself noted, implementing this generalized estimator-correlator structure would not be a simple task.

Actually, the importance of these results relates to the insights they give into optimal receiver structures. They show that the estimator-correlator interpretation of detection can be generalized to many different of signal and noise models. In other words, much of the

intuition provided by an estimator-correlator interpretation of the solution to the Gaussian signal in Gaussian noise detection problem carries over to more general situations.

The optimum and adaptive array processing problems have received a great deal of attention for nearly three decades, and consequently, there is an abundance of literature in these areas. It is beyond the scope of this dissertation to present a detailed review of the work already accomplished in these fields; however, we shall briefly point out several references of special interest.

Van Trees studied optimal array processing techniques in a classic report published two decades ago [2]. In particular, he examined the Gaussian signal in Gaussian noise detection problem, and derived several equivalent forms of the optimal structure, including the estimator-correlator. Other landmark papers were written by McDonough [11], Bryn [12], Edelblute and his colleagues [13], and Cox [14]. Adaptive array processing has been discussed in Monzingo and Miller [15], Haykin [16], Hudson [17], and Widrow and Stearns [18]. An extensive bibliography of current work in this area is contained in a new book edited by Sibul [19].

1.5 Overview of the Dissertation

Chapter 2 introduces the concepts, definitions, notation, and theorems that will be used in subsequent chapters. In particular, the Karhunen-Loève expansion is defined, and a method of calculating its basis from an arbitrary orthonormal basis is introduced.

In Chapter 3, the channel modeling problem is studied. Our approach is based on matrix representations of bounded, linear opera-

tors. The results have interesting implications in several disciplines, including the theory of non-stationary stochastic processes.

Solutions for $\underline{Q}(\cdot,\cdot)$ and $\underline{G}(\cdot,\cdot)$ are derived in Chapter 4. Spectral representations for $\underline{R}_N(\cdot,\cdot)$, $\underline{R}_y(\cdot,\cdot)$, and $\underline{R}_1(\cdot,\cdot)$ are used to solve the processor equations. Indeed, we will argue that this approach is not only a useful theoretical tool, but is also the key to adaptive implementation. The relationship between $f(\cdot)$ and $\underline{R}_y(\cdot,\cdot)$ is established and incorporated into the estimator-correlator structure.

It turns out that the relationship between $f(\cdot)$ and $\underline{R}_y(\cdot,\cdot)$ which is developed in Chapter 4 is not particularly convenient from an implementation point of view. Therefore, in Chapter 5, the stochastic identification problem is examined in further detail. Simultaneous diagonalization of the input and output covariance kernels is used to obtain a simplified relationship.

The numerical issues associated with adaptive implementation are studied in Chapter 6. Our results establish several interesting connections among canonical matrix decompositions and the Karhunen-Loève expansions of $\underline{G}(\cdot,\cdot)$ and $\underline{Q}(\cdot,\cdot)$. In addition, the CS decomposition is introduced as a numerically stable method for obtaining $\underline{G}(\cdot,\cdot)$ from array measurements.

In Chapter 7, numerical results are presented and evaluated. We consider an example of considerable practical interest; namely, detection in multipath propagation channels. Finally, conclusions and recommendations for further research are given in Chapter 8.

Chapter 2

MATHEMATICAL BACKGROUND

2.1 Introduction and Overview

This chapter introduces the concepts, definitions, notation, and theorems that will be used in subsequent chapters. Results will be stated, but proofs and derivations will be omitted. The interested reader can find careful developments of Hilbert space theory and linear operator theory in one of several standard texts [20] [21] [22].

We begin by establishing the mathematical structure in which signals are represented as elements in separable Hilbert spaces. This framework allows us to represent signal processing operations as bounded, linear operators defined over a Hilbert space of interest. Next, the problem of representing deterministic and stochastic signals is examined. We introduce generalized Fourier series expansions for both deterministic and stochastic signals, and the Karhunen-Loève expansion is defined. A new method of calculating the Karhunen-Loève basis is introduced. Finally, we discuss deterministic and stochastic operators, and demonstrate their usefulness in the context of this work.

2.2 Hilbert Spaces

2.2.1 Definition

A Hilbert space is defined as a complete inner product space [20]. An inner product space is a linear vector space endowed with a functional which maps the product space H x H onto the set of complex scalars. The functional is called the inner product, and will be

denoted by $< \cdot,\cdot >$. This inner product induces a norm onto H which is given by the definition

$$||x|| = < x,x >^{1/2} \qquad (2.2\text{-}1)$$

for all x in H.

Inner product spaces are complete if every Cauchy sequence converges to a point in the space.

Only separable Hilbert spaces shall be studied in this dissertation. A Hilbert space is separable only if an orthonormal basis exists in the space.

2.2.2 Examples

The set of all N-tuples of complex scalars is a Hilbert space with inner product

$$< x,y > = \sum_{i=1}^{N} x_i \, y_i^* \qquad (2.2\text{-}2)$$

for all x and y in H. This elementary Hilbert space is very useful in signal processing applications for representing finite-dimensional deterministic signals.

The space of all square-integrable functions defined on an interval (a, b) of the real line can be shown to be a separable Hilbert space [20]. In signal processing, this could represent all possible finite-energy waveforms received over a finite time interval T or spatial aperture A. In these cases, a suitable inner product is

$$\langle x, y \rangle = \int_T x(t)^* \, v(t) \, dt \qquad (2.2\text{-}3)$$

The norm $||x||$ represents the energy of a signal x.

If a received signal is more appropriately represented as a stochastic process, a useful inner product definition is:

$$\langle x,v \rangle = \int_T E\{x(t, \omega) \, v^*(t, \omega)\} \, dt \qquad (2.2\text{-}4)$$

Various combinations of purely temporal, purely spatial, deterministic, or stochastic signals can be chosen, and clearly, a large number of special cases can be studied in a common framework.

2.3 Deterministic Signal Representations

2.3.1 Fourier Series

Next, consider the problem of obtaining numerical representations for signals belonging to a separable Hilbert space H. Since we are only considering these Hilbert spaces, it is possible to find an orthonormal basis $\{\phi_k\}$ such that every element of H has the representation

$$x = \sum_k x_k \, \phi_k \qquad (2.3\text{-}1)$$

By definition, the orthonormal system $\{\phi_k\}$ is called an orthonormal basis for H, and each complex scalar x_k is called a Fourier coefficient of x. The Fourier coefficients are directly related to x through a simple inner product operation:

$$x_k = \langle x, \phi_k \rangle$$

The sequence of Fourier coefficients $\{x_k\}$ is called the representation of x with respect to $\{\phi_k\}$.

2.3.2 Interpretation of Generalized Fourier Series Expansions

Orthonormal expansions of deterministic and stochastic signals provide the fundamental connection between analog signals and their digital representations. For example, if $s(\cdot)$ is a deterministic signal bandlimited to the interval $(-\Omega, \Omega)$ on the angular frequency axis, then $s(\cdot)$ can be expanded in terms of a generalized Fourier series

$$s(t) = \sum_{n=-\infty}^{\infty} s(nT) \frac{\sin \Omega(t - nT)}{\Omega(t - nT)} \qquad (2.3\text{-}2)$$

The Fourier coefficients are samples of $s(\cdot)$. Frequently, it is often more convenient to work with the samples $\{s(nT)\}$ than the actual signal $s(\cdot)$. Moreover, it can be shown that L_2 is unitarily equivalent [20] with the space ℓ_2 of square-summable infinite sequences. This isomorphism, combined with the direct connection between a signal and its equivalent ℓ_2 representation, is the theoretical justification for sampled data or discrete-time signal processing. Problems formulated in terms of continuous time or continuous aperture measurements may be easier to solve in their equivalent ℓ_2 representations.

The choice of orthonormal basis depends on the class of signals the designer expects to encounter. Examples of bases frequently used in

signal processing applications are complex exponentials, prolate spheroidal functions, Karhunen-Loève eigenfunctions, and the standard basis. A judicious basis selection can often simplify the overall processor structure. A Hilbert space approach allows us to study signal processing systems and their mathematical representations from a fundamental and unified perspective.

2.4 Representing Stochastic Processes

2.4.1 Introduction

Until now we have been concerned with developing representations for deterministic signals. However, since the data measured at the array are produced by physical mechanisms which are inherently random in nature, deterministic descriptions of these phenomena are usually inadequate. Therefore, more general models of these processes based on probabilistic notions are essential.

Developing models for stochastic processes is necessarily different from those derived for their deterministic counterparts. However, it turns out that the Hilbert space concepts such as Fourier series, norms, and inner product are equally useful for characterizing stochastic as well as deterministic processes.

2.4.2 Definitions

A stochastic process is a generalization of the concept of a random variable [23] [24]. Let Ω be a fundamental sample space, and T denote an index set of an independent variable such as time or position. Then a stochastic process $x(\cdot,\cdot)$ is defined on the product space $T \times \Omega$ such

that for any $t \in T$, $x(t, \cdot)$ is a random variable. A stochastic vector process is an N-dimensional vector whose components are stochastic processes; that is

$$\underline{x}(\cdot,\cdot) = [x_1(\cdot,\cdot) \ \ldots \ x_N(\cdot,\cdot)]^T \qquad (2.4\text{-}1)$$

where $x_i(\cdot,\cdot)$ is a stochastic process for $i = 1,2,\ldots, N$. A complex-valued stochastic process $z(\cdot,\cdot)$ is a special case of a two-dimensional stochastic vector process:

$$z(t, \omega) = x(t, \omega) + jy(t, \omega) \qquad (2.4\text{-}2)$$

for all $t \in T$ and $\omega \in \Omega$.

Two particular forms of the index set T are important. If T is a sequence $\{t_1, t_2, t_3,\ldots\}$, then $x(\cdot,\cdot)$ is called a discrete-time stochastic process. On the other hand, if T is an interval of the real line, $x(\cdot,\cdot)$ is called a continuous-time stochastic process. In this dissertation, we will have the occasion to use both types of stochastic processes.

2.4.1 Canonical Representations and Karhunen-Loève Expansions

The rigorous definition of a stochastic process is actually quite abstract. It is an uncountable collection of measurable functions defined on a given probability space. A representation more suitable for numerical calculations is needed.

Any stochastic process $x(\cdot,\cdot)$ defined on the product space $T \times \Omega$ with continuous, square-integrable covariance kernel $R_X(\cdot,\cdot)$ can be expanded into a generalized Fourier series

$$x(t, \omega) = \sum_{i=1}^{\infty} x_i(\omega)\, \phi_i(t) \qquad (2.4\text{-}3)$$

where $\phi_i(\cdot)$ is an orthonormal basis, and

$$x_i(\omega) = \langle x, \phi_i \rangle$$

for all i [22]. The functional $\langle \cdot, \cdot \rangle$ is the standard inner product defined over the space of interest. The series coefficients $\{x_i(\cdot)\}$ are random variables whose numerical values depend on the particular realization $\omega \in \Omega$.

Equation (2.4-3) can be thought of as a decomposition, in which a function of two variables is expressed as a sum of products of functions of one variable. These decompositions are familiar from the theory of partial differential equations, and their significance in this context is essentially the same. Since the basis functions $\{\phi_i(\cdot)\}$ are deterministic, we can replace the study of an uncountable set of stochastic processes $\{x(\cdot,\cdot)\}$ by a countable set of random variables $\{x_i(\cdot)\}$. Moreover, the ideas from Hilbert space theory can be applied to this representation provided the inner product functional is chosen appropriately.

Equation (2.4-3) is a generalized Fourier series, hence any complete orthonormal basis $\{\phi_i(\cdot)\}$ spanning $L_2(a, b)$ can be used. However, one series is important enough to warrant a name. This representation is called the Karhunen-Loève orthonormal expansion [25] [26] of $x(\cdot,\cdot)$, and its usefulness can be explained as follows. If the unique Karhunen-

Loève basis is used, a nice "double orthogonality" results. The basis functions are orthonormal:

$$\int_T \phi_i(t)\, \phi_j^*(t)\, dt = \delta_{ij} \tag{2.4-4}$$

and so are the expansion coefficients:

$$E\{x_i(\omega)\, x_j^*(\omega)\} = \lambda_i^{(x)}\, \delta_{ij} \tag{2.4-5}$$

The basis providing for properties (2.4-4) and (2.4-5) can be found by solving the integral equation

$$\lambda_i^{(x)}\, \phi_i(t) = \int R_x(t, u)\, \phi_i(u)\, du \tag{2.4-6}$$

where $R_x(t, u) = E\{x(t, \omega)\, x^*(u, \omega)\}$ is the covariance function of process $x(\cdot,\cdot)$.

Another important result is Mercer's Theorem [22], which states that any positive definite, Hermitian, square-integrable kernel $k(\cdot,\cdot)$ can be expanded in a series representation:

$$k(t,s) = \sum_j \lambda_j^{(k)}\, \phi_j(t)\, \phi_j^*(s) \tag{2.4-7}$$

where $\{\lambda_j^{(k)}\}$ and $\{\phi_j(\cdot)\}$ are the eigenvalues and eigenfunctions of $k(\cdot,\cdot)$. This is the spectral representation for $k(\cdot,\cdot)$, one which is particularly convenient for numerical calculations. Since covariance

kernels meet the preceding conditions, they can be expanded in this form.

2.5 Calculating the Karhunen-Loève Expansion

2.5.1 Derivation

The nice properties of the Karhunen-Loève representation can be exploited to simplify the structure of the optimal processor. However, the computational problem involved with solving Equation (2.4-6) for the basis functions is difficult, and limits the practical applications of this approach. Another means of calculating the Karhunen-Loève expansion from arbitrary Fourier series representations would be very useful.

Let us demonstrate how this can be accomplished. Although the following derivation is for scalar stochastic processes, the results can easily be generalized to the vector case.

Suppose a random process $x(\cdot,\cdot)$ is expanded in terms of an arbitrary orthonormal basis:

$$x(t,\omega) = \sum_{i=1}^{\infty} x_i(\omega)\, \phi_i(t) \qquad (2.5\text{-}1)$$

The coefficients $\{x_i(\omega)\}$ are not necessarily orthogonal:

$$E\{x_i(\omega)\, x_j^*(\omega)\} = r_{ij} \qquad (2.5\text{-}2)$$

Next, write the covariance kernel $R_x(\cdot,\cdot)$ in terms of Equation (2.4-3):

$$R_x(t,u) = E\{x(t,\omega)\, x^*(u,\omega)\}$$

$$= E\{ \sum_{i=1}^{\infty} x_i(\omega)\, \Phi_i(t) \left[\sum_{j=1}^{\infty} x_j(\omega)\, \Phi_j(u) \right]^* \}$$

$$= E\{ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} x_i^*(\omega)\, x_j(\omega)\, \Phi_i^*(t)\, \Phi_j(t) \} \qquad (2.5\text{-}3)$$

Since the Fourier series converges to $x(\cdot,\cdot)$ in the mean-square sense, the expectation operator and the double sums can be interchanged [25]:

$$R_x(t,u) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} E\{x_i(\omega)\, x_j^*(\omega)\} \Phi_i(t)\, \Phi_j^*(u)$$

$$= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} r_{ij}\, \Phi_i(t)\, \Phi_j^*(u) \qquad (2.5\text{-}4)$$

The double sum is conviently written in vector-matrix notation:

$$R_x(t,u) = \underline{\Phi}^H(t)\, \underline{R}_x\, \underline{\Phi}(u) \qquad (2.5\text{-}5)$$

where

$$\underline{\Phi}(t) = [\Phi_1(t)\ \Phi_2(t)\ \ldots\ ]^T$$

and $\underline{R}_x$ is the infinite matrix of correlations. It is easy to show that $\underline{R}_x$ is Hermitian, which implies that $\underline{R}_x$ has a unique eigenvalue decomposition [22]:

$$\underline{R}_x = \underline{U}\, \underline{\Lambda}'\, \underline{U}^H \qquad (2.5\text{-}6)$$

Then $R_x(\cdot,\cdot)$ can be written

$$R_x(t,u) = \underline{\Phi}^H(t)\ \underline{R_x}\ \underline{\Phi}(u)$$

$$= \underline{\Phi}^H(t)\ \underline{U}\ \underline{\Lambda}'\ \underline{U}^H\ \underline{\Phi}(u)$$

$$= \sum_{k=1}^{\infty} \lambda_k^{(x)}\ \phi_k(t)\ \phi_k^*(u) \qquad (2.5\text{-}7)$$

The $\{\phi_k(\cdot)\}$ functions are a linear combination of the $\{\Phi_k(\cdot)\}$ basis:

$$\phi_k(t) = \sum_{\ell=1}^{\infty} u_{k\ell}^*\ \Phi_\ell(t) \qquad (2.5\text{-}8)$$

They are orthonormal:

$$\int_T \phi_i^*(t)\ \phi_j(t)\ dt = \int_T \Big[\sum_{m=1}^{\infty} u_{im}^*\ \Phi_m(t)\Big]\Big[\sum_{n=1}^{\infty} u_{jn}^*\ \Phi_n(t)\Big]^* dt$$

$$= \int_T \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} u_{im}^*\ u_{jn}\ \Phi_m(t)\ \Phi_n^*(t)\ dt$$

$$= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} u_{im}^*\ u_{jn} \int_T \Phi_m(t)\ \Phi_n^*(t)\ dt$$

$$= \sum_{m=1}^{\infty} u_{im}^*\ u_{jm} = \delta_{ij}$$

since the rows of $\underline{U}$ are orthogonal. Furthermore, the $\{\phi_k(\cdot)\}$ functions are complete, since each is a linear combination of a complete orthonormal basis. Therefore, the $\{\phi_k(\cdot)\}$ set is a complete orthonormal

basis, and since the basis which diagonalizes the matrix representation of $R_x(\cdot,\cdot)$ is unique, the functions are the Karhunen-Loève basis, the representation (2.5-7) is the spectral representation of $R_x(\cdot,\cdot)$, and the parameters $\{\lambda_k^{(x)}\}$ are the eigenvalues of integral equation (2.4-6).

In principle, this allows the Karhunen-Loève basis and expansion coefficients to be obtained from an arbitrary expansion, provided the second order statistics of its Fourier coefficients are known. The Karhunen-Loève basis is calculated with Equation (2.5-8), and the uncorrelated expansion coefficients are given by the transformation

$$\underline{x}_{KL}(\omega) = \underline{U}^H \underline{x}(\omega) \qquad (2.5\text{-}9)$$

where $\underline{x}_{KL}(\cdot)$ and $\underline{x}(\cdot)$ are column vectors containing the expansion coefficients. This transformation decorrelates the coefficients, since

$$E\{\underline{x}_{KL}(\omega)\ \underline{x}_{KL}^H(\omega)\} = \underline{U}^H\ E\{\underline{x}(\omega)\ \underline{x}^H(\omega)\}\underline{U} = \text{diag}\ [\lambda_1^{(x)}\ \lambda_2^{(x)}\ \ldots\ ]$$

In Chapter 6, we will show how to calculate the expansion directly from data with numerically stable algorithms.

2.5.2 Example

This example is based upon a measurement model used frequently by Nolte and his colleagues [27] [28]. Consider a time-limited stochastic process measured over the interval $(-T_o/2, T_o/2)$. Then $x(\cdot,\cdot)$ can be expanded into the Fourier series

$$x(t,\omega) = \sum_{n=-\infty}^{\infty} x_n(\omega)\ e^{j\omega_o nt} \qquad (2.5\text{-}10)$$

where

$$x_n(\omega) = \frac{1}{(2\pi T_o)1/2} \int_{T_o} x(t,\omega)e^{-j\omega_o nt}\ dt$$

In order to simplify the representation of $x(\cdot,\cdot)$, thev made the additional assumption that $x(\cdot,\cdot)$ is essentially bandlimited to $\pm\ L\omega_o$ radians/sec. Then

$$x(t,\omega) = \sum_{n=-L}^{L} x_n(\omega)e^{j\omega_o nt} \qquad (2.5\text{-}11)$$

Of course, the expansion coefficients $\{x_n(\cdot)\}$ are correlated. Can the Karhunen-Loève eigenfunctions and expansion coefficients be obtained from Equation (2.5-11)?

The covariance operator $R_x(\cdot,\cdot)$ is

$$R_x(t,u) = E\{x(t,\omega)\ x^*(u,\omega)\} = E\{\sum_{k=-L}^{L} x_k(\omega)\ \chi_k(t)\ [\sum_{\ell=-L}^{L} x_\ell^*(\omega)\ \chi_\ell^*(t)]\}$$

$$= \sum_{k=-L}^{L}\ \sum_{\ell=-L}^{L} E\{x_k(\omega)\ x_\ell^*(\omega)\}\ \chi_k(t)\ \chi_\ell^*(u)$$

$$= \sum_{k=-L}^{L}\ \sum_{\ell=-L}^{L} E\{x_k(\omega)\ x_\ell^*(\omega)\}\ e^{j\omega_o kt}\ e^{-j\omega_o \ell u} \qquad (2.5\text{-}12)$$

In vector-matrix notation, the Hermitian form (2.5-12) is given by

$$R_x(t,\omega) = \underline{X}^T(t)\ \underline{R}_{xx}\ \underline{X}^*(u) \qquad (2.5\text{-}13)$$

where

$$\underline{X}^T(t) = [e^{-j\omega_o Lt} \quad \ldots \quad e^{j\omega_o Lt}]$$

is a (2L + 1) x 1 column vector, and

$$\underline{R}_{xx} = E\{\underline{x}(\omega)\ \underline{x}^H(\omega)\} \qquad (2.5\text{-}14)$$

is the covariance matrix of the (2L + 1) Fourier coefficients. Performing an eigenvalue decomposition of $\underline{R}_{xx}$ gives

$$\underline{R}_{xx} = \underline{U}\ \underline{\Lambda}_x\ \underline{U}^H \qquad (2.5\text{-}15)$$

and substituting (2.5-15) into (2.5-13) leads to the result

$$R_x(t,u) = \underline{X}^T(t)\ \underline{U}\ \underline{\Lambda}_x\ \underline{U}^H\ \underline{X}^*(u) = \underline{\Phi}^T(t)\ \underline{\Lambda}_x\ \underline{\Phi}^*(u) \qquad (2.5\text{-}16)$$

It is easy to see that:

$$\underline{\Phi}(u) = \underline{U}^T\ \underline{X}(u) \qquad (2.5\text{-}17)$$

Rewriting (2.5-16) in terms of a sum yields:

$$R_x(t,u) = \sum_{n=1}^{2n+1} \lambda_n^{(x)}\ \phi_n(t)\ \phi_n^*(u) \qquad (2.5\text{-}18)$$

where $\phi_n(\cdot)$ is the nth element of the column vector $\underline{\Phi}(\cdot)$, and moreover,

$$\phi_n(t) = \underline{U}_n^T \underline{X}(t)$$

is a linear combination of the complex exponential terms. Equation (2.5-18) is the result; namely, it is the spectral representation of $R_x(\cdot,\cdot)$. Since the spectral representation is unique, the functions $\{\phi_n(\cdot)\}$ are the Karhunen-Loève basis, and the Karhunen-Loève expansion of $x(\cdot,\cdot)$ is

$$x(t,\omega) = \sum_{n=1}^{2N+1} x_{KL}(\omega)\ \phi_n(t)$$

The $\{x_{KL}(\cdot)\}$ coefficients can be obtained directly from (2.5-10) by a linear transformation:

$$\underline{x}_{KL}(\omega) = \underline{U}\ \underline{x}(\omega)$$

where

$$\underline{x}_{KL}(\omega) = [x_1(\omega) \ . \ . \ . \ x_{2N+1}(\omega)]^T$$

and

$$\underline{x}(\omega) = [x_L(\omega) \ . \ . \ . \ x_{-L}(\omega)]^T$$

The covariance matrix $\underline{R}_{xx}$ can be estimated from measured data by averaging over several sets of measurements. This procedure will be discussed in Chapter 6.

2.6 Deterministic and Stochastic Operators

2.6.1 Deterministic Operators

Given Hilbert spaces H_1 and H_2. A function $L(\cdot)$ which maps H_1 into H_2 is called a linear operator if, for all x and y in H_1, and complex scalars α:

$$L(x + y) = L(x) + L(y)$$

$$L(\alpha x) = \alpha L(x)$$

For convenience, we will write Lx instead of L(x). We will often write Lx(t) or $Lx(t,\omega)$ for x belonging to L_2 (deterministic signals) or $L_2 \times \Omega$ (stochastic signals) respectively. Since systems of practical interest are stable in the bounded-input, bounded-output sense [29], we need only be concerned with bounded linear operators. Mathematical representations for bounded linear operators shall be discussed in the next chapter.

2.6.2 Stochastic Operators and Their Representations

Clearly, deterministic descriptions of many signal transformations encountered in practical problems are inadequate. For example, wave propagation through realistic scattering media is described by partial differential equations whose coefficients are stochastic processes. Obviously, deterministic operator representations of these transformations are insufficient, and more general representations based on probabilistic notations are required.

A stochastic operator $\mathfrak{L}(\cdot)$ is a mapping defined over a Hilbert space H which is indexed with respect to a variable ω belonging to a

probability space (Ω, F, P) [4]. An example of a linear stochastic operator defined over C^N is an N x N matrix whose elements are random variables. ω could represent the value of an unknown physical quantity, such as the velocity of wave propagation in a medium.

A useful representation for stochastic operators defined over L_2 is the random Green's function representation. This description can be justified on physical grounds since one can demonstrate that solutions to the stochastic wave equation are written in this form [30] [31]. Furthermore, models for range spread, Doppler spread, and double spread channels can also be expressed as random Green's functions [1].

A random Green's function representation for $\mathfrak{f}(\cdot)$ is given by:

$$y(t, \omega) = \mathfrak{f}x(t, \omega) = \int_{-\infty}^{\infty} h(t, \tau, \omega)\, x(\tau, \omega)\, d\tau \qquad (2.6\text{-}1)$$

The random Green's function $h(\cdot,\cdot,\cdot)$ is the impulse response function for a linear stochastic system, and is a generalization of the impulse response function for a linear, time-variant deterministic system. This description of the medium establishes a connection between the physical description of realistic scattering media and the system theoretic representations familiar to engineers and applied mathematicians. A more detailed discussion of this approach to linear stochastic system modeling is contained in a recent monograph by Adomian [30].

2.7 Conclusions

We have established the mathematical structure which will be used throughout the dissertation. Deterministic and stochastic processes will be represented by generalized Fourier series expansions. Signal processing operations and stochastic transmission media will be modeled by bounded, linear operators defined over Hilbert spaces of deterministic and random signals.

The Karhunen-Loève expansion was defined and discussed in some detail. Its properties will be exploited later when the implementation problem is studied. This expansion can be obtained from arbitrary expansions by linear transformations of the basis functions and expansion coefficients.

Chapter 3

MATRIX REPRESENTATIONS OF PROPAGATION AND SCATTERING OPERATORS

3.1 Introduction

We shall begin by examining the channel modeling problem in detail. In this chapter, the goal is to derive mathematical representations for those deterministic and random transformations frequently encountered in practical array processing problems. We are especially interested in developing numerical representations that are convenient for digital signal processing applications, and that can be incorporated naturally into the estimator-correlator structure. The approach is based on matrix representations of bounded, linear operators.

This chapter contains several new results. The matrix representations for deterministic operators defined over the Hilbert space of bandlimited signals were introduced by Sibul and the author [32]. Furthermore, the results have interesting ramifications in the theory of non-stationary stochastic processes.

We obtain novel representations for vector-valued deterministic and stochastic signals, and for operators defined over these Hilbert spaces. Matrix representations for beamforming and multipath propagation channels are derived. The model for the multipath channel is a generalization of a model which recently appeared in the literature [33].

Section 3.2 introduces the fundamental ideas which shall be used throughout the chapter. In the next section, we introduce matrix representations for bounded linear operators, and show how they are useful

for discrete-time processing applications. To demonstrate their usefulness, matrix representations for delay, time stretching/compression, and simultaneous time delay and stretching given band limited signals will be worked out.

Next, matrix representations for stochastic operators will be developed. The representations are equivalent to those derived for deterministic transformations, and can be used to model spread scattering media.

In Section 3.4, matrix representations for vector-valued processes will be studied in some detail. It will be seen that convenient orthonormal bases can be constructed from simpler functions. Matrix representations for operators defined over these spaces are constructed. Actually, in general, the representations for these operators must be regarded as tensors, since they are indexed with respect to four indicies. Fortunately, if the basis is chosen judiciously, the operator representations can be greatly simplified.

Finally, matrix representations of multipath channels are derived. The result is a generalization of a model which recently appeared in the literature.

3.2 Deterministic Operators

3.2.1 Elementary Transformations

Let us define several elementary signal processing transformations, beginning with propagation delay:

$$r(t) = s(t - \tau), \quad \tau > 0$$

In the absence of noise, the received signal is a delaved version of the transmitted signal $s(\cdot)$.

Time stretching/compression is a generalization of Doppler shifing which applies to wide band signals. It is defined by

$$r(t) = s(\alpha t)$$

for $\alpha > 0$. Simultaneous stretching and delay is

$$r(t) = s(\alpha(t - \tau))$$

a combination of the two previous transformations.

These ideas can be generalized in a slightly more abstract setting. If $s(\cdot)$ represents a finite energy signal, then $s(\cdot)$ and $r(\cdot)$ can be represented as elements in L_2 or an appropriate subspace of L_2, which implies that the transformations defined above can be represented as linear operators mapping elements from an input space into an output space. The operators are defined implicitly:

$$r(t) = A_\tau s(t) = s(t - \tau)$$

The stretching/compression operator $A_\alpha(\cdot)$ is:

$$A_\alpha s(t) = s(\alpha t)$$

Simultaneous stretching and delay is

$$A_s\ s(t) = s(\alpha(t - \tau))$$

This operator is a cascade of $A_\alpha(\cdot)$ and $A_\tau(\cdot)$:

$$A_S(\cdot) = A_\alpha \, A_\tau(\cdot)$$

It is easy to show that all three operators are bounded and linear.

An explicit representation for a linear operator $\mathfrak{L}(\cdot)$ can be defined by a Green's function representation:

$$\mathfrak{L}s(t) = \int_{-\infty}^{\infty} h(t, \tau) \, x(\tau) \, d\tau$$

which is a convolution integral for a linear, time-variant system. The use of this representation in signal processing applications is well established [1] [34].

3.2.2 Matrix Representations of Bounded Linear Operators

The operators defined in the last section are useful in continuous time signal processing problems. However, for digital signal processing purposes, equivalent operators for sequences are needed. Furthermore, a representation which unifies the so-called implicit and explicit definitions presented above is highly desirable.

We propose using matrix representations of the continous time operators. The definition begins by expanding $r(\cdot)$ and $s(\cdot)$ into generalized Fourier series:

$$r(t) = \sum_k < r, \phi_k > \phi_k(t)$$

$$s(t) = \sum_k < s, \phi_k > \phi_k(t)$$

The coefficients $\{< r, \phi_k >\}$ and $\{< s, \phi_k >\}$ are the Fourier coefficients of $r(\cdot)$ and $s(\cdot)$ with respect to the basis $\{\phi_k\}$. Then the matrix representation $[\ell_{ij}]$ of an operator $\mathfrak{L}(\cdot)$, where

$$r(t) = \mathfrak{L}s(t)$$

is found by calculating the inner product

$$\ell_{ij} = < \mathfrak{L}\phi_j, \phi_i > \qquad (3.2\text{-}1)$$

The Fourier coefficients of r are obtained by evaluating the sum

$$r_i = \sum_j \ell_{ij} s_j$$

Of course, the model parameters $\{\ell_{ij}\}$ depend on the choice of basis functions used to model the signals. Suitable choices for practical applications include sampling functions, complex exponentials, prolate spheroidal functions, the standard basis, and Karhunen-Loève eigenfunctions. Some important properties of matrix representations are independent of the specific basis. This allows the modeling and identification problems to be examined from a fundamental and unified perspective.

3.2.3 Example: Time Delay

Let us work several specific examples to illustrate the preceeding ideas. Suppose that $s(\cdot)$ is a band limited signal whose Fourier transform vanishes outside the frequency interval $(-\sigma, \sigma)$. Then $s(\cdot)$ has the following generalized Fourier series representation:

$$s(t) = \sum_{n=-\infty}^{\infty} s(nT) \frac{\sin \sigma(t - nT)}{\sigma(t - nT)} \qquad (3.2\text{-}2)$$

where T is the sampling interval. If the sinc functions are normalized by the factor $1/\sqrt{T}$, they form an orthonormal basis for this subspace of L_2. The Fourier coefficients are simply the samples of $s(\cdot)$.

Recall that the delay operator was defined by the relation:

$$A_\tau s(t) = s(t - \tau)$$

If $s(\cdot)$ has representation (3.2-2), what is the matrix representation of $A_\tau(\cdot)$?

From the result presented in Equation (3.2-1):

$$a_{mn} = \langle A_\tau \phi_n, \phi_m \rangle = \frac{1}{T} \int_{-\infty}^{\infty} \frac{\sin \sigma(t - \tau - nT)}{\sigma(t - \tau - nT)} \frac{\sin \sigma(t - mT)}{\sigma(t - mT)} dt \qquad (3.2\text{-}3)$$

This integral is evaluated in the Appendix using contour integration. The result is:

$$a_{mn} = \frac{\sin \sigma(\tau - (m-n)T)}{\sigma(\tau - (m-n)T)} \qquad (3.2\text{-}4)$$

To check if the answer is reasonable, suppose the delay τ is an integer multiple of the sampling period T. Then

$$a_{mn} = \frac{\sin \sigma(kT - (m-n)T)}{\sigma(kT - (m-n)T)} = \frac{\sin \pi(k - (m-n))}{\pi(k - (m-n))}$$

which implies that

$$a_{mn} = \begin{cases} 0 & m-n \neq k \\ 1 & m-n = k \end{cases} \qquad (3.2\text{-}5)$$

and the matrix representation of $A_\tau(\cdot)$ reduces to a shift matrix analogous to the $\underline{Z}$ matrix introduced by Kailath and his colleagues [35].

3.2.4 Digression: Application to Non-stationary Random Processes

The matrix representation of delay operator $A_\tau(\cdot)$ has some interesting ramifications in the theory of non-stationary stochastic processes. In an attempt to measure "how far" an arbitrary discrete-time stochastic process deviates from wide sense stationarity, Kailath and his colleagues introduced the concept of displacement rank [36]. The displacement matrix $\delta\underline{R}$ is defined by:

$$\delta\underline{R} = \underline{R} - \underline{Z}\,\underline{R}\,\underline{Z}^T \qquad (3.2\text{-}6)$$

where $\underline{R}$ is the infinite-dimensional covariance matrix of the stochastic process $\{x(\cdot,\cdot)\}$, and $\underline{Z}$ is the lower shift matrix given by Equation (3.2-4) for $k = 1$. They argued that the rank of $\delta\underline{R}$ is a measure of how far $x(\cdot,\cdot)$ deviates from wide-sense stationarity.

The $\underline{Z}$ matrix can be interpreted as the matrix representation for $A_\tau(\cdot)$ with respect to the standard basis. Calculating the matrix elements in this case is trivial; however, only integer multiples of the sampling interval can be represented. Our representations allow these ideas to be generalized to arbitrary time shifts.

3.2.5 Time Compression/Expansion

Next, the matrix representation for a time stretch/compression operator will be calculated. Let $A_\alpha(\cdot)$ be defined as:

$$A_\alpha s(t) = s(\alpha t)$$

for $\alpha > 0$.

Its matrix representation is given by

$$a_{mn} = \frac{1}{T} \int_{-\infty}^{\infty} \frac{\sin \sigma(\alpha t - nT)}{\sigma(\alpha t - nT)} \frac{\sin \sigma(t - mT)}{\sigma(t - mT)} \, dt$$

The calculation is straightforward and similar to evaluating (3.2-3). The answer is

$$a_{mn} = \frac{\sin(\sigma T(\alpha m - n)/\alpha)}{(\sigma T(\alpha m - n)/\alpha)} \qquad (3.2\text{-}7)$$

If there is no stretching or compression, then $\alpha = 1$, and

$$a_{mn} = \frac{\sin \pi(m - n)}{\pi(m - n)} = \delta_{mn}$$

which is the identity operator.

3.2.6 Combined Propagation Delay and Stretching/Compression

This example illustrates several interesting ideas which relate cascaded operators to their matrix representations.

Begin by defining the operator $A_S(\cdot)$ implicitly by the relation

$$A_S s(t) = s(\alpha(t - \tau))$$

which is the combination of delay and stretching. We have already shown that $A_S(\cdot)$ is equivalent to cascaded $A_\alpha(\cdot)$ and $A_\tau(\cdot)$:

$$A_S(\cdot) = A_\alpha\, A_\tau(\cdot)$$

Its matrix representation can be calculated in two ways. The first is by the original definition:

$$a_{mn} = \frac{1}{T} \int_{-\infty}^{\infty} \frac{\sin \sigma(\alpha t - \alpha\tau - nT)}{\sigma(\alpha t - \alpha\tau - nT)} \frac{\sin \sigma(t - mT)}{\sigma(t - mT)}\, dt$$

with the result

$$a_{mn} = \frac{\sin \sigma(\alpha\tau - (\alpha m - n)T/\alpha)}{\sigma(\alpha\tau - (\alpha m - n)T/\alpha)} \qquad (3.2\text{-}8)$$

The second is by cascading the matrix representations of $A_S(\cdot)$ and $A_\tau(\cdot)$. This can be performed by multiplying the two infinite matrices together:

$$a_{mn}(\alpha,\tau) = \sum_{j=-\infty}^{\infty} a_{mj}(\tau)\, a_{jn}(\alpha) \qquad (3.2\text{-}9)$$

where the $\{a_{mj}(\cdot)\}$ are the delay operator matrix elements, and the $\{a_{jn}(\cdot)\}$ are the stretching/compression elements. From Equations (3.2-3) and (3.2-7), we want to calculate

$$\sum_{j=-\infty}^{\infty} \frac{\sin \sigma(\tau - (m - j)T)}{\sigma(\tau - (m - j)T)} \cdot \frac{\sin(\sigma T(\alpha j - n)/\alpha)}{(\sigma T(\alpha j - n)/\alpha)} \qquad (3.2\text{-}10)$$

This sum can be evaluated by rearranging Equation (3.2-10) and applying Shannon's Sampling Theorem. Rearranging (3.2-10) yields

$$\sum_{j=-\infty}^{\infty} \frac{\sin \sigma T(j - n/\alpha)}{\sigma T(j - n/\alpha)} \cdot \frac{\sin(\sigma(\tau - (m - j)T)}{\sigma(\tau - (m - j)T)} \qquad (3.2\text{-}11)$$

Let $k = m - j$. Then the sum becomes

$$\sum_{k=\infty}^{-\infty} \frac{\sin \sigma T(m - k - n/\alpha)}{\sigma T(m - k - n/\alpha)} \frac{\sin \sigma(\tau - kT)}{\sigma(\tau - kT)}$$

$$= \sum_{k=-\infty}^{\infty} \frac{\sin \sigma(mT - kT - nT/\alpha)}{\sigma(mT - kT - nT/\alpha)} \frac{\sin \sigma(\tau - kT)}{\sigma(\tau - kT)} \qquad (3.2\text{-}12)$$

Now for bandlimited $f(\cdot)$,

$$f(t) = \sum_{k=-\infty}^{\infty} f(kT) \frac{\sin \sigma(t - kT)}{\sigma(t - kT)}$$

We immediately equate

$$f(kT) = \frac{\sin \sigma(mT - kT - nT/\alpha)}{\sigma(mT - kT - nT/\alpha)} = \frac{\sin \sigma(kT - mT + nT/\alpha)}{\sigma(kT - mT + nT/\alpha)} \qquad (3.2\text{-}13)$$

which is bandlimited to $(-\alpha\sigma, \alpha\sigma)$. Therefore, the right hand side of (3.2-12) is

$$\frac{\sin \sigma(\tau - mT - nT/\alpha)}{\sigma(\tau - mT - nT/\alpha)} = \frac{\sin \sigma(\alpha\tau - (\alpha m - n)T)/\alpha)}{\sigma(\alpha\tau - (\alpha m - n)T)/\alpha)}$$

which equals the result obtained by evaluating inner product functional directly.

This example demonstrates that cascading linear operators defined over L_2 spaces and their matrix representations is equivalent. Computationally speaking, it is usually easier to determine the matrix representation of cascaded operators directly from the definition rather than by multiplying large or infinite matrices. It is interesting to note that by applying Equation (3.2-9) to cascaded operators, we obtain a method for computing products of infinite matrices.

3.3 Matrix Representations of Stochastic Operators

3.3.1 General

Two fundamentally different matrix representations are associated with stochastic operators: Those characterizing $\mathfrak{L}(\cdot)$ itself, and those representing its statistical properties. Both are useful in signal and array processing applications. In this section, we shall examine only the first type of representations. Those defined for statistical measures of $\mathfrak{L}(\cdot)$ will be examined in Section 3.3.3.

The matrix representation of a stochastic operator $\mathfrak{L}(\cdot)$ with respect to a basis $\{\phi_k\}$ is defined exactly as in the deterministic case:

$$\ell_{ij}(\omega) = \langle \mathfrak{L}\phi_j, \phi_i \rangle \qquad (3.3\text{-}1)$$

Each matrix element $\{\ell_{ij}(\cdot)\}$ is a random variable whose numerical value depends upon the particular realization ω of $\mathfrak{L}(\cdot)$.

3.3.2 Examples

Suppose $\mathfrak{L}(\cdot)$ is defined in terms of a random Green's function representation:

$$\mathfrak{L}(\cdot) = \int_{-\infty}^{\infty} h(t, \tau, \omega)(\cdot)\, d\tau$$

The matrix representation of $\mathfrak{L}(\cdot)$ is:

$$\ell_{ij}(\omega) = \langle \mathfrak{L}\phi_j, \phi_i \rangle = \int_{-\infty}^{\infty} \{ \int_{-\infty}^{\infty} h(t, \tau, \omega)\, \phi_j(\tau)\, d\tau \}\, \phi_i^*(t)\, dt$$

$$= \int_{-\infty}^{\infty} \int h(t, \tau, \omega)\, \phi_j(\tau)\, \phi_i^*(t)\, d\tau\, dt \qquad (3.3\text{-}2)$$

Notice that the time dependence of $h(\cdot,\cdot,\cdot)$ does not appear in its matrix representation. Therefore, we have a means of representing a time-varying system with coefficients that are time-invariant.

The random Green's function representation can be used to model range spread and double spread media. For example, a signal $y(\cdot,\cdot)$ returned from a range-spread channel can be expressed

$$y(t, \omega) = \int_{L} b_R(\lambda, \omega)\, s(t - \lambda)\, d\lambda$$

where $b_R(\cdot,\cdot)$ is a sample function from a zero mean complex Gaussian random process, and λ is the spatial variable. Its matrix representation is

$$\ell_{ij}(\omega) = \langle \mathfrak{L}\phi_j, \phi_i \rangle$$

$$= \int_{-\infty}^{\infty} \{ \int_{L} b_R(\lambda, \omega)\, \phi_j(t - \lambda)\, d\lambda \}\, \phi_i^*(t)\, dt$$

$$= \int_{-\infty}^{\infty} \int_{L} b_R(\lambda, \omega)\, \phi_j(t - \lambda)\, \phi_i^*(t)\, d\lambda\, dt$$

This result can be simplified by exchanging the order of integration. This is permitted provided $b_R(\cdot,\cdot)$ is continuous for all ω. Then

$$\ell_{ij}(\omega) = \int_{L} b_R(\lambda, \omega) \left\{ \int_{-\infty}^{\infty} \phi_j(t - \lambda)\, \phi_i^*(t)\, dt \right\} d\lambda$$

But the integral within brackets is the definition for the matrix representation of a deterministic time delay operator $A_\lambda(\cdot)$:

$$< A_\lambda\, \phi_j,\, \phi_i > = \int_{-\infty}^{\infty} \phi_j(t - \lambda)\, \phi_i^*(t)\, dt$$

Therefore,

$$\ell_{ij}(\omega) = \int_{L} b_R(\lambda, \omega)\, a_{ij}(\lambda)\, d\lambda \qquad (3.3\text{-}4)$$

In a similar fashion, the return from a double spread medium (spread with respect to time and frequency) can be represented by

$$y(t, \omega) = \int_{-\infty}^{\infty} b(t - \lambda/2, \lambda, \omega) s(t - \lambda)\, d\lambda$$

and its matrix representation is:

$$\ell_{ij}(\omega) = < \mathfrak{L}\phi_j,\, \phi_i >$$

$$= \int_{-\infty}^{\infty} \int b(t - \lambda/2, \lambda, \omega)\, \phi_j(t - \lambda)\, \phi_i^*(t)\, d\lambda\, dt \qquad (3.3\text{-}4)$$

3.3.3 Matrix Representations of Covariance Kernels

We will find that the matrix representations of integral covariance kernels provide useful insights into several problems which will be studied in this work. Their derivation is straightforward, since it can be shown that if $\{\phi_k(\cdot)\}$ is a complete orthonormal basis spanning $L_2(T)$, then the functions

$$\Phi_{k\ell}(t, s) = \phi_k(t)\, \phi_\ell^*(s)$$

are a complete orthonormal basis spanning the product space $L_2(T)$ x $L_2(T)$ [22]. This is the space of all square-integrable functions $k(\cdot,\cdot)$, and since we have assumed all stochastic processes have square-integrable covariance kernels, they belong to this space. Therefore, in terms of the basis $\{\Phi_{k\ell}(\cdot,\cdot)\}$,

$$R_x(t, s) = \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} \langle R_x, \phi_k\, \phi_\ell^* \rangle\, \phi_k(t)\, \phi_\ell^*(s) \qquad (3.3\text{-}5)$$

where

$$\langle R_x, \phi_k\, \phi_\ell^* \rangle = \int_T \int_T R_x(t, s)\, \phi_k^*(t)\, \phi_\ell(s)\, ds\, dt \qquad (3.3\text{-}6)$$

Equations (3.3-5) and (3.3-6) can be expressed in terms of expansion coefficient cross-correlations, because

$$R_x(t,s) = E\{x(t, \omega)\, x^*(s, \omega)\}$$

By exchanging integration order in Equation (3.3-6),

$$< R_x, \phi_k\, \phi_\ell^* > = E\{\int_T \int_T x(t, \omega)\, x^*(s, \omega)\, \phi_k^*(t)\, \phi_\ell(s)\, ds\, dt\}$$

$$= E\{[\int_T x(t, \omega)\, \phi_k^*(t)\, dt][\int_T x^*(s, \omega)\, \phi_\ell(s)\, ds]\}$$

$$= E\{x_k(\omega)\, x_\ell^*(\omega)\}$$

Therefore, $R_x(\cdot,\cdot)$ can be written in the form

$$R_x(t, s) = \underline{\Phi}^H(s)\, \underline{R}_x\, \underline{\Phi}(t) \qquad (3.3\text{-}7)$$

where

$$\underline{\Phi}^T(t) = [\phi_1(t)\ \phi_2(t)\ \ldots\ ]$$

and the infinite matrix $\underline{R}_x$ is a matrix of Fourier coefficient cross-correlations. $\underline{R}_x$ is the matrix representation of $R_x(\cdot,\cdot)$ with respect to $\{\phi_k(\cdot)\}$. Notice that when the Karhunen-Loève basis is used the off-diagonal terms of $\underline{R}_x$ vanish. In other words, the matrix representation of $R_x(\cdot,\cdot)$ is diagonalized.

3.4 Representations for Vector-Valued Processes and Transformations

3.4.1 Signal Representations

Numerical representations of vector-valued signals are essential for array processing applications. Therefore, we next take up the

problem of representing signals $\underline{x}(\cdot)$ defined over the product space $L_2(a, b) \times C^N$, where the interval (a, b) is the index set of continuous time variable t, and C^N is the set of complex-valued N-tuples representing N channels of data. We seek equivalent ℓ_2 representations of this data.

The first step in solving the problem is defining a suitable inner product functional over the Hilbert space $L_2(a, b) \times C^N$. One such functional is

$$< \underline{x}, \underline{v} > = \int_a^b \underline{x}^T(t)\, \underline{v}^*(t)\, dt \qquad (3.4\text{-}1)$$

which is convenient for signal processing applications, because the induced norm $||\underline{x}||$ has the interpretation as total signal energy summed over all N channels.

Provided a complete orthonormal basis exists,

$$\underline{x}(t) = \sum_{i=1}^{\infty} x_i\, \underline{\phi}_i(t) \qquad (3.4\text{-}2)$$

and as before, the Fourier coefficients are given by

$$x_i = < \underline{x}, \underline{\phi}_i > \qquad (3.4\text{-}3)$$

for all i.

In principle, Equations (3.4-2) and (3.4-3) can be used to obtain a numerical representation for $\underline{x}(\cdot)$. However, it is not readily apparent how a suitable basis $\{\underline{\phi}_k(\cdot)\}$ spanning the product space $L_2(a, b) \times C^N$ can be found. A method of constructing an orthonormal basis from those

which span the individual spaces $L_2(a, b)$ and C^N is needed. The following theorem is a generalization of a result discussed in Gohberg and Goldberg [22], and allows an appropriate orthonormal basis to be constructed.

Theorem III.1: If $\{\phi_i\}$ is an orthonormal basis for $L_2(a, b)$, and $\{\underline{u}_j\}$ is an orthonormal basis for C^N, then the functions

$$\underline{\Phi}_{ij}(t) = \phi_i(t)\, \underline{u}_j \tag{3.4-4}$$

form an orthonormal basis for $L_2(a, b) \times C^N$, for $i = 1, 2, \ldots$, and $j = 1, 2, \ldots, N$.

Proof: To demonstrate that $\{\underline{\Phi}_{ij}(\cdot)\}$ is an orthonormal basis, it suffices to show that any two basis elements are orthonormal, and that $< \underline{g}, \underline{\Phi}_{ij} > = 0$ implies that $\underline{g}$ is identically zero almost everywhere.

Orthonormality follows easily:

$$< \underline{\Phi}_{k\ell}, \underline{\Phi}_{mn} > = \int_a^b \underline{\phi}_{k\ell}^T(t)\, \underline{\phi}_{mn}^*(t)\, dt$$

$$\int_a^b \phi_k(t)\, \underline{e}_\ell^T\, \underline{e}_m^*\, \phi_n^*(t)\, dt = \underline{e}_\ell^T\, \underline{e}_m^* \int_a^b \phi_k(t)\, \phi_n^*(t)\, dt = \delta_{\ell m}\, \delta_{kn}$$

Next, it is necessary to show that if $\underline{g} \in L_2(a, b) \times C^N$ and if

$$< \underline{g}, \underline{\Phi}_{ij} > = 0 \tag{3.4-5}$$

then $\underline{g} = \underline{0}$ almost everywhere.

Expanding Equation (3.4-5) gives

$$< \underline{g}, \underline{\Phi}_{ij} > = \int_a^b \underline{g}^T(t)\, \phi_i^*(t)\, \underline{e}_j^*\, dt = \underline{g}_i^T\, \underline{e}_j^* = 0 \qquad (3.4\text{-}6)$$

where $\underline{g}_i$ is defined by

$$\underline{g}_i = \left[\int_a^b g_1(t)\, \phi_i^*(t)\, dt \; . \; . \; . \int_a^b g_N(t)\, \phi_i^*(t)\, dt \right]^T \qquad (3.4\text{-}7)$$

an N-tuple of complex scalars. Now, since $\{\underline{e}_j\}$ is an orthonormal basis for C^N, Equation (3.4-6) implies $\underline{g}_i = \underline{0}$ for all i. This means that

$$\int_a^b g_m(t)\, \phi_i^*(t) = 0 \qquad (3.4\text{-}8)$$

where (3.4-8) is the m^{th} element of $\underline{g}_i$. (3.4-8) immediately implies that for m = 1, 2,..., N, $g_m(\cdot)$ is identically zero almost everywhere, because $\{\phi_i(\cdot)\}$ is an orthonormal basis for $L_2(a,b)$. It follows from Equations (3.4-6) and (3.4-8) that for (3.4-5) to hold for all possible i and j, $\underline{g}$ is identically $\underline{0}$ almost everywhere. Therefore, from Theorem 11.3 in Gohberg and Goldberg, $\{\underline{\Phi}_{ij}(\cdot)\}$ is an orthonormal basis for $L_2(a, b) \times C^N$.

Now that an orthonormal basis for $L_2(a, b) \times C^N$ has been constructed, a representation for elements $\underline{x}(\cdot)$ belonging to it can be developed:

$$\underline{x}(t) = \sum_{i=1}^{\infty} \sum_{j=1}^{N} x_{ij}\, \phi_i(t)\, \underline{e}_j$$

where

$$x_{ij} = \langle \underline{x}, \underline{\Phi}_{ij} \rangle$$

are the generalized Fourier coefficients. Since the norm of $\underline{x}(\cdot)$ is finite, by Parseval's equality

$$\sum_{i=1}^{\infty} \sum_{j=1}^{N} |x_{ij}|^2 < \infty$$

from which we conclude the $\{x_{ij}\}$ coefficients belong to ℓ_2.

3.4.2 Matrix Representations for Linear Operators Over $L_2(a, b) \times C^N$

Processing operations such as beamsteering or propagation channel effects such as angular spreading can be interpreted as linear operators mapping elements from $L_2(a, b) \times C^N$ into itself. Therefore, for the same reasons as before, we seek matrix representations for bounded, linear operators defined over this space.

Let $L(\cdot)$ be a bounded linear operator on $L_2(a, b) \times C^N$, and define the element $\underline{y}(\cdot)$ by

$$\underline{y}(t) = L\underline{x}(t)$$

for $\underline{x} \in H$. Then

$$L\underline{x}(t) = \sum_{k=1}^{\infty} \sum_{\ell=1}^{N} \langle \underline{x}, \underline{\Phi}_{k\ell} \rangle L\underline{\Phi}_{k\ell}(t) \qquad (3.4\text{-}9)$$

But

$$L\underline{\Phi}_{k\ell}(t) = \sum_{i=1}^{\infty} \sum_{j=1}^{N} < L\underline{\Phi}_{k\ell}, \underline{\Phi}_{ij} > \underline{\Phi}_{ij}(t) \qquad (3.4\text{-}10)$$

Combining Equations (3.4-9) and (3.4-10) yields

$$L\underline{x}(t) = \sum_{k} \sum_{\ell} < \underline{x}, \underline{\Phi}_{k\ell} > \left[\sum_{i} \sum_{j} < L\underline{\Phi}_{k\ell}, \underline{\Phi}_{ij} > \underline{\Phi}_{ij}(t)\right]$$

$$= \sum_{i} \sum_{j} \sum_{k} \sum_{\ell} < L\underline{\Phi}_{k\ell}, \underline{\Phi}_{ij} > < \underline{x}, \Phi_{k\ell} > \underline{\Phi}_{ij}(t) \qquad (3.4\text{-}11)$$

Now, this equation expresses $\underline{y}(\cdot)$ in terms of $\{\underline{\Phi}_{ij}(\cdot)\}$, and of course,

$$\underline{y}(t) = \sum_{i} \sum_{j} y_{ij} \underline{\Phi}_{ij}(t) = \sum_{i} \sum_{j} < \underline{y}, \underline{\Phi}_{ij} > \underline{\Phi}_{ij}(t) \qquad (3.4\text{-}12)$$

Equating coefficients $< \underline{y}, \underline{\Phi}_{ij} >$ in (3.4-11) to the equivalent representation in Equation (3.4-12) implies that

$$< \underline{y}, \underline{\Phi}_{ij} > = \sum_{k} \sum_{\ell} < L\underline{\Phi}_{k\ell}, \underline{\Phi}_{ij} > < \underline{x}, \underline{\Phi}_{k\ell} >$$

or that

$$y_{ij} = \sum_{k} \sum_{\ell} \ell_{ijk\ell} x_{k\ell} \qquad (3.4\text{-}13)$$

where

$$\ell_{ijk\ell} = < L\underline{\Phi}_{k\ell}, \underline{\Phi}_{ij} >$$

for all i and k, and for j and ℓ = 1, 2,..., N. This result which relates the representation of $\underline{x}(\cdot)$ to the representation of $\underline{y}(\cdot)$ leads to the following definition:

Let $\{\underline{\Phi}_{ij}(\cdot)\}$ be an orthonormal basis for $L_2(a, b) \times C^N$, and let $L(\cdot)$ be a bounded linear operator defined over H. The matrix representation of $L(\cdot)$ with respect to basis $\{\underline{\Phi}_{ij}(\cdot)\}$ is given by

$$\ell_{ijk\ell} = \langle L\underline{\Phi}_{k\ell}, \underline{\Phi}_{ij} \rangle \qquad (3.4\text{-}14)$$

for all i and k, and for ℓ and j = 1, 2,..., N.

Representations for stochastic operators defined over $L_2(a, b) \times C^N$ are obtained the same way:

$$\ell_{ijk\ell}(\omega) = \langle \mathfrak{L}\underline{\Phi}_{k\ell}, \underline{\Phi}_{ij} \rangle = \langle \mathfrak{L}\phi_k \underline{e}_\ell, \phi_i \underline{e}_j \rangle \qquad (3.4\text{-}15)$$

The $\{\ell_{ijk\ell}(\cdot)\}$ parameters relate the generalized Fourier coefficients of the input signal to those of the output process analogous to Equation (3.4-13).

Strictly speaking, the representation $\{\ell_{ijk\ell}\}$ must be regarded as a tensor rather than a two-dimensional matrix, since the parameters are indexed with respect to four variables. Obviously, this is not the most numerically convenient representation one might want to use in an application. However, the representation can be simplified by choosing the basis properly.

3.4.3 Beamsteering Operator

The example we will work in this section is to find a matrix representation for a beamsteering operator $L_B(\cdot)$:

$$L_B\, \underline{s}(t) = [s(t - \tau) \ \ldots \ s(t - (m - 1)\tau)]^T$$

$L_B(\cdot)$ represents the processing needed to steer the response of a uniformly spaced line array towards a desired look direction. The vector $\underline{s}(\cdot)$ is defined by

$$\underline{s}(t) = [s(t) \ . \ . \ . \ s(t)]^T \tag{3.4-16}$$

for $-\pi \leq t \leq \pi$. Assuming $s(\cdot)$ is periodic with period $T = 2\pi$,

$$s(t) = \sum_{n=-\infty}^{\infty} s_n\, e^{jnt} \tag{3.4-17}$$

where the fundamental frequency ω_o is unity.

Convenient representations for $\underline{s}(\cdot)$ and $\underline{y}(\cdot)$ are

$$\underline{y}(t) = \sum_{i=-\infty}^{\infty} \sum_{j=1}^{M} y_{ij}\, \phi_i(t)\, \underline{e}_j \tag{3.4-18}$$

$$\underline{s}(t) = \sum_{i=-\infty}^{\infty} \sum_{j=1}^{M} s_{ij}\, \phi_i(t)\, \underline{e}_j \tag{3.4-19}$$

The matrix representation of $L_B(\cdot)$ relates the $\{s_{ij}\}$ coefficients to the $\{y_{ij}\}$ coefficients

$$y_{ij} = \sum_{k=-\infty}^{\infty} \sum_{\ell=1}^{N} \ell_{ijk\ell}\, s_{k\ell} \tag{3.4-20}$$

and the coefficients $\{\ell_{k\ell mn}\}$ are found by

$$\ell_{k\ell mn} = \langle L_B \, \underline{\Phi}_{mn}, \, \underline{\Phi}_{k\ell} \rangle = \langle L_B \, \phi_m \, \underline{e}_n, \, \phi_k \, \underline{e}_\ell \rangle$$

The product space $L_2(-\pi, \pi) \times C^M$ is a Hilbert space with inner product functional

$$\langle \underline{x}, \underline{v} \rangle = \int_{-\pi}^{\pi} \underline{x}^T(t) \, \underline{y}^*(t) \, dt \qquad (3.4\text{-}21)$$

for all elements $\underline{x}$ and $\underline{y}$ in H.

Let us demonstrate how a judicious choice of basis simplifies the matrix representation of $L_B(\cdot)$. A natural selection for the $\{\phi_k(\cdot)\}$ functions is the complex exponentials:

$$\phi_k(t) = e^{jkt}$$

and the standard basis for the $\{\underline{e}_k\}$ vectors ($k = 1, \ldots, M$). Then

$$\ell_{k\ell mn} = \int_{-\pi}^{\pi} L_{\tau_\ell} \phi_k(t) \, \phi_m^*(t) \, dt \cdot \underline{e}_\ell^T \underline{e}_n = \delta_{\ell n} \int_{-\pi}^{\pi} e^{jk(t - \tau_\ell)} \, e^{-jnt} \, dt$$

$$\delta_{\ell n} \, e^{-jk\tau_\ell} \int_{-\pi}^{\pi} e^{j(k - m)t} \, dt = e^{-jk\tau_\ell} \, \delta_{km} \, \delta_{\ell n} \qquad (3.4\text{-}22)$$

for $k = \ldots, -1, 0, 1, \ldots$, and $\ell = 1, 2, \ldots, M$. The τ_ℓ is the delay introduced into the ℓth sensor measurement.

Equation (3.4-22) shows how the selection of basis simplifies the matrix representation. Its kth column is

$$[e^{-jk\tau_1} \ldots e^{-jk\tau_m}]^T$$

which is the steering vector for a narrowband signal centered at $\omega = k\omega_0$. This shows that a narrowband steering vector is simply a special case of a more general beamsteering operator which applies to wideband processing.

3.4.4 Matrix Representations for Multipath Channels

In the next example, matrix representations for multipath propagation channels will be derived. The received data are defined over a Hilbert space $L_2(a, b) \times C^M$ and have the canonical representation

$$\underline{r}(t, \omega) = \sum_{i=-\infty}^{\infty} \sum_{j=1}^{M} r_{ij}(\omega)\, \phi_i(t)\, \underline{e}_j$$

The transmitted signal $s(\cdot)$ will be modeled as a deterministic, scalar, periodic function with Fourier series expansion

$$s(t) = \sum_{k=-\infty}^{\infty} s_k\, e^{jk\omega_0 t}$$

How can the representation $\{s_k\}$ be related to the $\{r_{ij}(\cdot)\}$ parameters in a multipath channel?

In a multipath environment, energy from a common source arrives at the array from several directions. A model must take three phenomena into account; namely, a time delay $\tau_p(\cdot)$ common to every path due to range delay, a random real-valued scalar representing propagation loss due to distance, and time delay due to wavefront propagation across the

array itself. The latter time delay depends on the arrival direction of the wavefront and the array geometry.

Without loss of generality, the matrix representation will be derived for a single arriving wavefront. The result is easily generalized to multiple wavefronts due to linearity of the medium.

In the absence of noise, the array measurement at time t is:

$$\underline{r}(t, \omega) = \begin{vmatrix} b(\omega)\ s(t - \tau_p - \tau_1) \\ \vdots \\ b(\omega)\ s(t - \tau_p - \tau_M) \end{vmatrix} \qquad (3.4\text{-}23)$$

where $b(\cdot)$ represents propagation loss, τ_p is the delay due to distance traveled, and τ_i is the time delay at the ith sensor due to the array geometry.

Rewriting Equation (3.4-23) in operator notation gives

$$\underline{r}(t, \omega) = \underline{f}\underline{s}(t)$$

where

$$\underline{f}\underline{s}(t) = \begin{vmatrix} L_{\tau_p} & & 0 \\ & \ddots & \\ 0 & & L_{\tau_p} \end{vmatrix} \begin{vmatrix} b(\omega) & & 0 \\ & \ddots & \\ 0 & & b(\omega) \end{vmatrix} \begin{vmatrix} L_{\tau_1} & & 0 \\ & \ddots & \\ 0 & & L_{\tau_M} \end{vmatrix} \begin{vmatrix} s(t) \\ \vdots \\ s(t) \end{vmatrix}$$

and both $\underline{r}(\cdot,\cdot)$ and $\underline{s}(\cdot)$ can be expanded as:

$$\underline{r}(t, \omega) = \sum_{i=-\infty}^{\infty} \sum_{j=1}^{M} r_{ij}(\omega)\ \phi_i(t)\ \underline{e}_j$$

$$\underline{s}(t) = \sum_{i=-\infty}^{\infty} \sum_{j=1}^{M} s_{ij}\, \phi_i(t)\, \underline{e}_j$$

The matrix representation of $\mathfrak{L}(\cdot)$ is computed by

$$\ell_{k\ell mn}(\omega) = < \mathfrak{L}\underline{\Phi}_{mn},\ \underline{\Phi}_{k\ell} >$$

where the inner product functional $< \cdot, \cdot >$ is

$$< \underline{x},\ \underline{y} > = \int_T \underline{x}^T(t)\ \underline{y}^*(t)\ dt$$

Before proceeding with the calculation, a suitable basis must be chosen. We have already proved that a basis spanning $L_2(a, b) \times C^M$ can be constructed from products of basis functions spanning the individual spaces. In principle, one could choose any arbitrary $\{\phi_k\}$ and $\{\underline{e}_k\}$ from a wide range of choices. However, a judicious selection of basis might simplify the matrix representation. Since $s(\cdot)$ is periodic, the natural choice for the scalar basis is the set of complex exponentials. It is not clear what constitutes a good choice for the $\{\underline{e}_k\}$ basis, therefore, let us try the standard basis first.

In terms of the standard basis,

$$\ell_{k\ell mn}(\omega) = < \mathfrak{L}\underline{\Phi}_{mn},\ \underline{\Phi}_{k\ell} > = \int_{T_0} (\mathfrak{L}\phi_m(t)\ \underline{e}_n)^T\ \phi_k^*(t)\ \underline{e}_\ell^*\ dt \qquad (3.4\text{-}24)$$

where T_0 is the observation interval, and $T_0 = 2\pi/\omega_0$.

Now,

$$\mathfrak{L}\phi_m(t)\,\underline{e}_n = \begin{bmatrix} L_{\tau_1} & & 0 \\ & \ddots & \\ 0 & & L_{\tau_M} \end{bmatrix} \begin{bmatrix} b(\omega) & & 0 \\ & \ddots & \\ 0 & & b(\omega) \end{bmatrix} \begin{bmatrix} L_{\tau_p} & & 0 \\ & \ddots & \\ 0 & & L_{\tau_p} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \phi_m(t) \\ 0 \end{bmatrix}$$

$$= [0 \ \ldots \ 0 \ \ L_{\tau_n} \ b(\omega) \ L_{\tau_p} \ \phi_m(t)]^T$$

and

$$L_{\tau_n} \ b(\omega) \ L_{\tau_p} \ \phi_m(t) = \frac{1}{(2\pi T_0)^{1/2}} \ b(\omega) \ \exp\{jm\omega_0(t - \tau_n - \tau_p)\}$$

$$= b(\omega) \ \exp\{-jm\omega_0\tau_p\} \ \exp\{-jm\omega_0\tau_n\} \ \phi_m(t) \qquad (3.4\text{-}25)$$

Substituting (3.4-25) into Equation (3.4-24) gives

$$\ell_{k\ell mn}(\omega) = \delta_{\ell n} \ \exp\{-jm\omega_0\tau_p\} \ \exp\{-jm\omega_0\tau_n\} \ b(\omega) \int_{T_0} \phi_m(t) \ \phi_k^*(t) \ dt$$

$$= \delta_{mk} \ \delta_{\ell n} \ b(\omega) \ e^{-jm\omega_0\tau_p} \qquad (3.4\text{-}26)$$

for m or k integer, and ℓ or n = 1, 2,..., M. The coefficients vanish if $m \neq k$ or $\ell \neq n$, which means that the tensor representation of $\mathfrak{L}(\cdot)$ reduces to a simpler matrix structure. Indeed, the kth column of the matrix is

$$b(\omega)\ e^{-jk\omega_0\tau_p} \begin{bmatrix} e^{-jk\omega_0\tau_1} \\ \vdots \\ e^{-jk\omega_0\tau_M} \end{bmatrix} \qquad (3.4\text{-}27)$$

The column vector is simply a steering vector for a narrowband signal of frequency $k\omega_0$, meaning that the matrix operator can be interpreted as an infinite set of steering vectors acting on the individual frequency components.

In order to generalize the result to multipath channels, the representation for each path is summed element wise, which is permitted since we have assumed a linear medium. The k-th column of the matrix becomes

$$\sum_{i=1}^{p} b_i(\omega)\ e^{-jk\omega_0\tau_{p_i}}\ \underline{v}_{ki} \qquad (3.4\text{-}28)$$

where p is the number of paths, τ_{p_i} is the delay due to the i-th path, and $\{\underline{v}_{ki}\}$ are the narrowband steering vectors associated with the k-th Fourier coefficient of $s(\cdot)$ and the i-th wavefront arriving at the array.

3.4.5 Discussion

A special case of matrix representation (3.4-28) has appeared in the literature. Recently, Paulraj and Kailath [33] developed a multipath channel model in the context of an optimal beamforming problem.

Their model for the array output $\underline{r}(\cdot,\cdot)$ in the absence of noise is written:

$$\underline{r}(t, \omega) = \underline{A}\ \underline{f}(\omega)\ s(t) \qquad (3.4\text{-}29)$$

where $s(\cdot)$ is a deterministic, narrowband signal, $\underline{A}$ is an M x p matrix of steering vectors representing the arrival directions of p wavefronts, and $\underline{f}(\cdot)$ is a p x 1 random vector whose ith element is a complex scalar representing the path loss and phase shift in the ith path.

Let us demonstrate that their multipath channel model is simply a special case of our model. Assuming that $s(\cdot)$ is a narrowband signal in the form

$$s(t) = e^{j\omega_o t}$$

then $s(\cdot)$ is a special case of the Fourier expansion used in the last section,

$$s_k = \begin{cases} 0 & k \neq 1 \\ 1 & k = 1 \end{cases}$$

and rearranging $\underline{A}\ \underline{f}(\omega)$ gives

$$\underline{A}\ \underline{f}(\omega) = \sum_{i=1}^{p} f_i(\omega)\ \underline{a}_i$$

From Equation (3.4-28), the matrix representation relating the series expansion coefficients of $s(\cdot)$ to those of $\underline{r}(\cdot,\cdot)$ is

$$\sum_{i=1}^{p} b_i(\omega)\, e^{-j\omega_o \tau_p}\, \underline{v}_i$$

Equating the $\{f_i(\cdot)\}$ terms with the $\{b_i(\cdot)\exp(-j\omega_o\tau_p)\}$ coefficients, and equating the steering vectors $\{\underline{a}_i\}$ with $\{\underline{v}_i\}$ demonstrates that Paulraj and Kailath's multipath model reduces to a special case of Equation (3.4-28).

3.5 Conclusions

Matrix representations of bounded, linear operators provide a convenient means of modeling deterministic and stochastic signal transformations which arise in array processing problems. They are well suited for digital signal processing applications since they are based upon orthonormal representations of those signals under examination. It is clear that selecting a basis judiciously can simplify the operator modeling problem a great deal.

We have obtained new results that shed new light into classical signal theory issues. In particular, the representations for time delay operators gives new insight into the theory of non-stationary stochastic processes. Of course, our interest in matrix representations goes beyond theoretical considerations alone. The final purpose of this work is to develop stochastic and deterministic models that can easily be incorporated into the estimator-correlator for subsequent identification in conjunction with detection.

These comments lead to a very important issue which deserves further discussion. All of the models presented in this chapter are

based on orthonormal series expansions of deterministic and random processes. Of course, many other methods of signal and system modeling have been extensively studied over the last forty years, especially state space and parametric or rational transfer function models. Why has a Fourier series expansion approach to the modeling problem been selected?

First, this approach gives considerable insight into the nature and structure of $\mathcal{L}(\cdot)$. For example, we showed how a beamforming operator for wideband, periodic signals can be represented as a matrix whose columns can be interpreted as narrowband steering vectors. This is an intuitively pleasing result, and it gives physical meaning to an abstract result that might be completely missed if other models were used.

Second, these representations allow for a great deal of flexibility. Any basis spanning the signal space can be used, and we have already demonstrated that a judicious selection of basis can simplify the modeling problem. Furthermore, this allows the detection problem to be solved generically for an arbitrary set of generalized Fourier coefficients. This is important, since discrete Fourier transforms are used frequently in practical applications.

Matrix operators avoid difficulties inherent with representing non-stationary signals and systems. Recall that matrix representations of time-varying systems are themselves time-invariant. This property simplifies the modeling and identification problems considerably. On the other hand, modeling time-varying, stochastic systems with rational

transfer function representations is exceedingly difficult, and in our opinion, not mathematically mature.

The most important reason why this approach has been selected relates to the structure of the estimator kernel $\underline{G}(\cdot,\cdot)$. Orthonormal series expansions will be used to implement $\underline{G}(\cdot,\cdot)$, and the matrix representations for $\underline{f}(\cdot)$ can be incorporated into the structure in a straightforward manner. The Karhunen-Loève basis is preferred because it simplifies the solution for $\underline{G}(\cdot,\cdot)$, and it represents a fundamental approach to the stochastic system modeling and identification problem.

Chapter 4

IMPLEMENTING THE OPTIMAL STRUCTURE

4.1 Introduction

This chapter addresses several crucial issues which arise in the process of solving the estimator-correlator equations:

$$\int_T \underline{R}_N(t, u)\, \underline{Q}(u, z)\, du = \delta(t - z)\underline{I} \qquad (4.1\text{-}1)$$

$$\int_T \underline{R}_1(t, u)\, \underline{G}(u, z)\, du = \underline{R}_V(t, z) \qquad (4.1\text{-}2)$$

First, it is necessary to establish a suitable mathematical representation for the covariance and filter kernels in order to solve Equations (4.1-1) and (4.1-2). It is not clear from the definitions of $\underline{R}_1(\cdot,\cdot)$, $\underline{R}_V(\cdot,\cdot)$, and $\underline{R}_N(\cdot,\cdot)$ alone how to formulate solutions for $\underline{Q}(\cdot,\cdot)$ and $\underline{G}(\cdot,\cdot)$. Moreover, obtaining numerical representations which are reasonably easy to manipulate in hardware is essential for adaptive realization of the processor.

Another issue which must be resolved is more fundamental. So far, we have presupposed perfect a priori knowledge of the covariance kernels. However, this is an unreasonable assumption in most practical situations; therefore, it is necessary to estimate the kernels from array measurements. Kernels $\underline{R}_1(\cdot,\cdot)$ and $\underline{R}_N(\cdot,\cdot)$ can be estimated directly from array data using standard adaptive techniques. On the other hand, $\underline{R}_V(\cdot,\cdot)$ can not be estimated directly from array data, since the signal component is always obscured by additive measurement noise.

If a priori knowledge of $\underline{R}_y(\cdot,\cdot)$ is unknown or incomplete, a means of estimating it in conjunction with detection must be found. Incomplete knowledge of $\underline{R}v(\cdot,\cdot)$ represents a fundamental limitation which must be overcome if the optimal structure is to be realized.

Third, the identification scheme for $\underline{R}_y(\cdot,\cdot)$ must be signal independent. In other words, once the channel operator $\underline{\ell}(\cdot)$ has been identified based on probing signal measurements, we should be able to use the results to estimate $\underline{R}_y(\cdot,\cdot)$ for an arbitrary transmitted signal.

Finally, a means of incorporating a priori knowledge of the scattering channel into the detector structure is needed. For example, we may know that $\underline{\ell}(\cdot)$ belongs to a certain class of channels characterized by one or more parameters. When one or more of these parameters are known, or if bounds can be placed on their values, the ability to incorporate this information into the processor would be useful.

We begin by reviewing the Karhunen-Loève orthonormal expansion and its relationship to the spectral representation of linear operators. This expansion is the fundamental tool which will be used to analyze and implement the processor. We assert that it provides the theoretical means to solve for the processor, gives considerable insight into its mathematical structure, and establishes a link between theoretical analysis and implementation.

Next, we will show how these ideas can be used to solve for the processor structure. Finding $\underline{Q}(\cdot,\cdot)$ is straightforward. Solving for the estimator branch is somewhat more difficult, but much more

interesting. Once again, a Karhunèn-Loeve representation provides the solution for $\underline{G}(\cdot,\cdot)$, and also suggests an implementation scheme. Furthermore, it provides useful insights into the modeling and identification problem. In Section 4.6, the relationship between $\mathfrak{L}(\cdot)$ and the estimator kernel is derived. Identification of $\mathfrak{L}(\cdot)$ is required in order to calculate the conditional mean estimate.

4.2 Covariance Kernel Representations

The key to solving Equations (4.1-1) and (4.1-2) is the spectral representation of covariance kernels $\underline{R}_N(\cdot,\cdot)$, $\underline{R}_y(\cdot,\cdot)$, and $\underline{R}_1(\cdot,\cdot)$.

The relationship between the Karhunen-Loève expansion of a process $\underline{x}(\cdot,\cdot)$ and the spectral representation of its covariance kernel $\underline{R}_x(\cdot,\cdot)$ can be easily demonstrated. Recall that $\underline{R}_x(\cdot,\cdot)$ is defined by:

$$\underline{R}_x(t, u) = E\{\underline{x}(t, \omega)\ \underline{x}^H(u, \omega)\} \tag{4.2-1}$$

for t and u within the observation interval T. Expressing Equation (4.2-1) in terms of the Karhunen-Loève expansions gives

$$\underline{R}_x(t, u) = E\Big\{\Big[\sum_{k=1}^{\infty} x_k(\omega)\ \underline{\phi}_k(t)\Big]\Big[\sum_{\ell=1}^{\infty} x_\ell(\omega)\ \underline{\phi}_\ell(u)\Big]^H\Big\}$$

$$= E\Big\{\sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} x_k(\omega)\ x_\ell^*(\omega)\ \underline{\phi}_k(t)\ \underline{\phi}_\ell^H(u)\Big\}$$

$$= \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} E\{x_k(\omega)\ x_\ell^*(\omega)\}\ \underline{\phi}_k(t)\ \underline{\phi}_\ell^H(u) \tag{4.2-2}$$

However, the expansion coefficients are uncorrelated, therefore, Equation (4.2-2) reduces to

$$\underline{R}_x(t, u) = \sum_{k=1}^{\infty} E\{|x_k(\omega)|^2\} \, \underline{\phi}_k(t) \, \underline{\phi}_k^H(u)$$

Furthermore, the $E\{|x_k(\omega)|^2\}$ terms correspond to the eigenvalues $\{\lambda_k^{(x)}\}$ of the integral equation whose kernel is $\underline{R}_x(\cdot,\cdot)$. Therefore:

$$\underline{R}_x(t,u) = \sum_{k=1}^{\infty} \lambda_k^{(x)} \, \underline{\phi}_k(t) \, \underline{\phi}_k^H(u) \qquad (4.2\text{-}3)$$

which is the spectral representation. This discussion shows how the uncorrelated expansion coefficients simplify the numerical representation of the covariance function, because they reduce a double sum representation to a single sum representation. Stated another way, the Karhunen-Loève basis diagonalizes the matrix representation of $\underline{R}_x(\cdot,\cdot)$.

4.3 The Inverse Filter Branch

4.3.1 Derivation

To illustrate how these representations are used, consider Equation (4.1-1):

$$\int_T \underline{R}_N(t, u) \, \underline{Q}(u, z) \, du = \delta(t - z)\underline{I}$$

and solve for the inverse filter $\underline{Q}(\cdot,\cdot)$.

From (4.2-3):

$$\underline{R}_N(t, u) = \sum_{k=1}^{\infty} \lambda_k^{(n)} \underline{\phi}_k(t) \underline{\phi}_k^H(u) \qquad (4.3\text{-}1)$$

We assert that the inverse operator $\underline{Q}(\cdot,\cdot)$ is

$$\underline{Q}(t, u) = \sum_{k=1}^{\infty} \frac{1}{\lambda_k^{(n)}} \underline{\phi}_k(t) \underline{\phi}_k^H(u) \qquad (4.3\text{-}2)$$

In order to prove this assertion, substitute Equations (4.3-1) and (4.3-2) into (4.1-1):

$$\int_T \left[\sum_{k=1}^{\infty} \lambda_k^{(n)} \underline{\phi}_k(t) \underline{\phi}_k^H(u) \right] \left[\sum_{\ell=1}^{\infty} \frac{1}{\lambda_\ell^{(n)}} \underline{\phi}_\ell(u) \underline{\phi}_\ell^H(z) \right] du \qquad (4.3\text{-}3)$$

Both sums converge uniformly; therefore, (4.3-3) can be rearranged as follows:

$$\int_T \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} \frac{\lambda_k^{(n)}}{\lambda_\ell^{(n)}} \underline{\phi}_k(t) \underline{\phi}_k^H(u) \underline{\phi}_\ell(u) \underline{\phi}_\ell^H(z) \, du$$

$$= \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} \frac{\lambda_k^{(n)}}{\lambda_\ell^{(n)}} \underline{\phi}_k(t) \int_T \underline{\phi}_k^H(u) \underline{\phi}_\ell(u) \, du \, \underline{\phi}_\ell^H(z) \qquad (4.3\text{-}4)$$

Since the basis functions are orthogonal, Equation (4.3-1) reduces to the form

$$\sum_{k=1}^{\infty} \underline{\phi}_k(t)\, \underline{\phi}_k^H(z) \qquad (4.3\text{-}5)$$

It can be shown that (4.3-5) equals $\delta(t - z)\underline{I}$, where $\delta(\cdot)$ is the Dirac delta function. In other words, $\underline{R}_N(\cdot,\cdot)$ and $\underline{Q}(\cdot,\cdot)$ are inverse kernels, and the inverse filter solution is given by (4.3-2).

4.3.2 Discussion

Equation (4.3-2) is the solution for $\underline{Q}(\cdot,\cdot)$ expressed in spectral form. Clearly, once the spectral representation of $\underline{R}_N(\cdot,\cdot)$ is found, calculating $\underline{Q}(\cdot,\cdot)$ is straightforward, because one only needs to calculate the reciprocals of the eigenvalues $\{\lambda_k^{(n)}\}$. This simple procedure demonstrates the usefulness of a spectral representation approach.

Although the preceding derivation for $\underline{Q}(\cdot,\cdot)$ is mathematically correct, several difficulties must be resolved before considering an implementation based on this solution. First, we have yet to discuss how the eigenvalues and eigenfunctions used to represent $\underline{R}_N(\cdot,\cdot)$ are found. Solving for them even given perfect knowledge of $\underline{R}_N(\cdot,\cdot)$ means solving a matrix integral equation, clearly not a simple task. Furthermore, since $\underline{R}_N(\cdot,\cdot)$ is usually unknown a priori, the kernel must be estimated from array data. Exactly how this can be done remains to be seen. Finally, the numerical issues involved with hardware implementation must be examined. Since the calculations associated with the preceeding discussion are often be performed on finite-precision fixed point hardware, robust computational algorithms ought to be used.

Each of these issues will be addressed later in the dissertation. However, for now let us assume that these difficulties can be resolved, and turn our attention to solving for the estimator kernel $\underline{G}(\cdot,\cdot)$.

4.4 The Estimator Branch

4.4.1 Derivation

In order to solve Equation (4.1-2), suppose that $\underline{R}_y(\cdot,\cdot)$ and $\underline{R}_1(\cdot,\cdot)$ can be expanded with respect to the same orthonormal basis $\{\underline{\phi}_k(\cdot)\}$:

$$\underline{R}_y(t, u) = \sum_{k=1}^{\infty} \lambda_k^{(y)} \underline{\phi}_k(t) \underline{\phi}_k^H(u) \qquad (4.4\text{-}1)$$

$$\underline{R}_1(t, u) = \sum_{k=1}^{\infty} \lambda_k^{(r)} \underline{\phi}_k(t) \underline{\phi}_k^H(u) \qquad (4.4\text{-}2)$$

The basis is not the same as used to expand $\underline{Q}(\cdot,\cdot)$. We claim that $\underline{G}(\cdot,\cdot)$ is given by

$$\underline{G}(t, u) = \sum_{k=1}^{\infty} \frac{\lambda_k^{(y)}}{\lambda_k^{(r)}} \underline{\phi}_k(t) \underline{\phi}_k^H(u) \qquad (4.4\text{-}3)$$

To prove this assertion, substitute (4.4-2) and (4.4-3) into the left hand side of Equation (4.1-2):

$$\int_T \Big[\sum_{k=1}^{\infty} \lambda_k^{(r)} \underline{\phi}_k(t) \underline{\phi}_k^H(u) \Big] \Big[\sum_{\ell=1}^{\infty} \frac{\lambda_\ell^{(y)}}{\lambda_\ell^{(r)}} \underline{\phi}_\ell(u) \underline{\phi}_\ell^H(z) \Big] \, du \qquad (4.4\text{-}4)$$

Both infinite series converge uniformly; therefore, (4.4-4) can be rearranged, yielding

$$\sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} \lambda_k^{(r)} \frac{\lambda_\ell^{(y)}}{\lambda_\ell^{(r)}} \underline{\phi}_k(t) \left\{ \int_T \underline{\phi}_k^H(u) \, \underline{\phi}_\ell(u) \, du \right\} \cdot \underline{\phi}_k^H(z) \qquad (4.4\text{-}5)$$

Once again appealing to the orthogonality condition of the basis functions leads to the result that (4.4-5) simplifies into

$$\sum_{k=1}^{\infty} \lambda_k^{(y)} \, \underline{\phi}_k(t) \, \underline{\phi}_k^H(z) = \underline{R}_y(t, z) \qquad (4.4\text{-}6)$$

which proves that Equation (4.4-3) is the solution for $\underline{G}(\cdot,\cdot)$.

4.4.2 Discussion

The solution for the estimator branch kernel depends on a crucial assumption; specifically, the ability to expand the covariance kernels $\underline{R}_y(\cdot,\cdot)$ and $\underline{R}_1(\cdot,\cdot)$ with respect to the same basis $\{\underline{\Phi}_k(\cdot)\}$. This is equivalent to simultaneously solving the equations

$$\lambda_k^{(x)} \, \underline{\Phi}_k(t) = \int_T \underline{R}_y(t, u) \, \underline{\Phi}_k(u) \, du \qquad (4.4\text{-}7)$$

$$\lambda_k^{(r)} \, \underline{\Phi}_k(t) = \int_T \underline{R}_1(t, u) \, \underline{\Phi}_k(u) \, du \qquad (4.4\text{-}8)$$

with the same set of orthogonal eigenfunctions $\{\underline{\Phi}_k(\cdot)\}$. From a more abstract point of view, this is equivalent to simultaneous diagonaliza-

tion of two linear operators [22]. It is essential to establish the conditions when this is possible.

It turns out that simultaneous diagonalization is trivial if the noise covariance kernel $\underline{R}_N(\cdot,\cdot)$ corresponds to the covariance function of a white noise process. Since

$$\underline{R}_1(\cdot,\cdot) = \underline{R}_y(\cdot,\cdot) + \underline{R}_N(\cdot,\cdot) \tag{4.4-9}$$

if

$$\underline{R}_N(t, u) = \delta(t - u)\underline{I} \tag{4.4-10}$$

then by substituting Equations (4.4-9) and (4.4-10) into Equation (4.4-8), it is clear that $\{\underline{\phi}_k(\cdot)\}$ solves both relations. In other words, if the noise process $\underline{n}(\cdot,\cdot)$ is spatially and temporally white, simultaneous diagonalization of $\underline{R}_1(\cdot,\cdot)$ and $\underline{R}_y(\cdot,\cdot)$ is straightforward.

However, making a white noise assumption as a part of the problem formulation is unrealistic. Therefore, it is necessary to establish the conditions when simultaneous diagonalization is possible given a colored noise process $\underline{n}(\cdot,\cdot)$.

Simultaneous diagonalization can be achieved by decorrelating the Fourier coefficients of $\underline{n}(\cdot,\cdot)$. The various relations among the covariance and filter operators is most easily seen with the ℓ_2 representations for the various signals and operators involved. In terms of Fourier series representations, the measurement model

$$\underline{r}(t,\omega) = \underline{f}\underline{s}(t) + \underline{n}(t,\omega)$$

is equivalent to

$$\underline{r}(\omega) = \underline{y}(\omega) + \underline{n}(\omega) \qquad (4.4\text{-}11)$$

where the coefficients $\underline{r}(\cdot)$, $\underline{y}(\cdot)$, and $\underline{n}(\cdot)$ are in column vector form. In terms of matrix representations, the relation

$$\underline{R}_I(\cdot,\cdot) = \underline{R}_y(\cdot,\cdot) + \underline{R}_N(\cdot,\cdot)$$

is equivalent to

$$\underline{R}_I = \underline{R}_y + \underline{R}_N \qquad (4.4\text{-}12)$$

Define a matrix filter $\underline{C}$ operating on $\underline{n}(\cdot)$ such that

$$\underline{n}' = \underline{C}\,\underline{n}$$

and

$$E\{\underline{n}'(\underline{n}')^H\} = \underline{I}$$

The structure of $\underline{C}$ depends on how the filter operator is implemented. For example, if Gram-Schmidt orthogonalization of the Fourier coefficient vector $\underline{n}(\cdot)$ is performed, $\underline{C}$ is an upper or lower triangular matrix.

The decorrelation can be performed by premultiplying $\underline{r}(\cdot)$ by $\underline{C}$:

$$\underline{C}\,\underline{r}(\omega) = \underline{C}\,\underline{y}(\omega) + \underline{C}\,\underline{n}(\omega)$$

Clearly, $\underline{C}$ alters the relation among $\underline{R}_I$, $\underline{R}_y$, and $\underline{R}_N$ in Equation

(4.4-12). The modified noise covariance kernel becomes an identity operator; however, it is not obvious that the modified signal plus noise and signal covariance kernels can be simultaneously diagonalized. Is this still possible?

The answer to this question is yes. We have

$$E\{\underline{r}'(\omega)(\underline{r}'(\omega))^H\} = E\{\underline{C}\ \underline{r}(\omega)(\underline{r}^H(\omega))\ \underline{C}^H\} = \underline{C}\ \underline{R}_1\ \underline{C}^H =$$

$$\underline{C}\ \underline{R}_y\ \underline{C}^H + \underline{C}\ \underline{R}_N\ \underline{C}^H = \underline{C}\ \underline{R}_y\ \underline{C}^H + \underline{I} \qquad (4.4\text{-}13)$$

or,

$$\underline{R}_1' = \underline{R}_y' + \underline{I} \qquad (4.4\text{-}14)$$

A theorem from functional analysis [22] states that a necessary and sufficient condition for simultaneously diagonalizing two linear operators $A(\cdot)$ and $B(\cdot)$ is that

$$AB(\cdot) = BA(\cdot) \qquad (4.4\text{-}15)$$

Now,

$$\underline{R}_1' = \underline{R}_y' + \underline{I} = \underline{R}_y'(\underline{I} + \underline{R}_y'^{-1})$$

$$= \underline{R}_y'(\underline{R}_y' + \underline{I})\ \underline{R}_y'^{-1} = \underline{R}_y'(\underline{R}_1')\ \underline{R}_y'^{-1}$$

Therefore,

$$\underline{R}_1' \underline{R}_y' = \underline{R}_y' \underline{R}_1' \underline{R}_y'^{-1} \underline{R}_y' = \underline{R}_y' \underline{R}_1'$$

which shows that the modified operators $\underline{R}_1'$ and $\underline{R}_y'$ are simultaneously diagonalizable. Moreover, since $\underline{R}_1'$ and $\underline{R}_y'$ are positive definite Hermitian forms, they can be diagonalized into spectral form with a unique unitary operator $\underline{U}'$. The operator $\underline{U}'$ diagonalizes both $\underline{R}_1'$ and $\underline{R}_y'$ simultaneously, since $\underline{R}_1'$ is the sum of $\underline{R}_y'$ and an identity operator. Therefore, in matrix form:

$$\underline{G}' = \underline{R}_y'(\underline{R}_1')^{-1} = \underline{U}' \underline{\Lambda}'_y \underline{U}'^H \underline{U}' \underline{\Lambda}_r'^{-1} \underline{U}'^H$$

$$= \underline{U}' \underline{\Lambda}_y' \underline{\Lambda}_r^{-1} \underline{U}'^H$$

$$= \underline{U}' \text{ diag}[\lambda_1^{(y')}/\lambda_1^{(r')} \ \ldots] \underline{U}'$$

The spectral form is obtained by decorrelating $\underline{n}(\cdot)$ when the noise process is colored.

4.5 The Structure of $G(\cdot,\cdot)$

The solution for $\underline{G}(\cdot,\cdot)$ is in the elegant form of a spectral representation parameterized by the eigenvalues of $\underline{R}_1(\cdot,\cdot)$ and $\underline{R}_y(\cdot,\cdot)$. But its significance goes beyond an abstract theoretical result of little or no practical interest. On the contrary, not only does this mathematical structure give considerable insight into the relationships among detection, estimation, modeling, and system identification, it is the

key to practical implementation as well. Why do we make these assertions?

First, evaluating the conditional mean estimate from array measurements is simple. This can be seen clearly by expanding $\underline{G}(\cdot,\cdot)$ in terms of Equation (4.4-3):

$$\int_T \underline{G}(t, u)\, \underline{r}(u, \omega)\, du = \int_T \left[\sum_k \frac{\lambda_k^{(y)}}{\lambda_k^{(r)}}\, \underline{\phi}_k(t)\, \underline{\phi}_k^H(u)\right] \underline{r}(u, \omega)\, du$$

$$= \sum_k \frac{\lambda_k^{(y)}}{\lambda_k^{(r)}} < \underline{r}, \underline{\phi}_k > \underline{\phi}_k(t) = \sum_k \frac{\lambda_k^{(y)}}{\lambda_k^{(r)}}\, r_k(\omega)\, \underline{\phi}_k(t)$$

The coefficients $\{r_k(\cdot)\}$ are obtained from the array measurements by an inner product operation. Calculating the conditional mean is straightforward because each coefficient is multiplied by a real number. This operation is analogous to postmultiplying an N x 1 vector of coefficients by an N x N diagonal matrix where only N multiples are needed. The preceding remarks suggest this structure reduces computational complexity.

The parameters $\{\lambda_k^{(r)}\}$ and $\{\lambda_k^{(y)}\}$ give insight into the channel operator identification problem, and suggest a method of obtaining $\underline{G}(\cdot,\cdot)$ adaptively. Recall that the $\{\lambda_k^{(r)}\}$ terms are the variances of the expansion coefficients of $\underline{r}(\cdot,\cdot)$ given H_1:

$$\lambda_k^{(r)} = E\{|r_k(\omega)|^2\} \qquad (4.5\text{-}1)$$

The $\{\lambda_k^{(y)}\}$ parameters are the variances of the signal process coefficients:

$$\lambda_k^{(y)} = E\{|v_k(\omega)|^2\} \qquad (4.5\text{-}2)$$

Therefore, the spectral representation of $\underline{G}(\cdot,\cdot)$ transforms the covariance kernel estimation problem into a parameter estimation problem. The significance of the parameters is obvious from Equations (4.5-1) and (4.5-2), and in addition, these definitions suggest how they can be obtained from array data. The $\{\lambda_k^{(r)}\}$ parameters can be estimated by performing the inner product operation $< \underline{r}, \phi_k >$, taking the magnitude squared of the result, and averaging over L independent measurements. On the other hand, obtaining the $\{\lambda_k^{(y)}\}$ parameters is not as straightforward, since uncorrupted measurements of the signal process $\underline{y}(\cdot,\cdot)$ are not available. These parameters represent a priori knowledge which is needed to calculate the conditional mean. If they are unknown, another estimation procedure must be found. This problem will be examined further in Chapter 6.

The other issue that must be addressed is how the parameters $\{\lambda_k^{(y)}\}$ relate to the channel operator $\mathfrak{L}(\cdot)$. This relationship will be studied in the next section.

4.6 Relating $\mathfrak{L}(\cdot)$ to the Unavailable Eigenvalues

The relationship between $\{\lambda_k^{(y)}\}$ parameters and $\mathfrak{L}(\cdot)$ is obtained in terms of the matrix representations for $\underline{R}_y(\cdot,\cdot)$ and $\mathfrak{L}(\cdot)$:

$$\underline{R}_y = E\{\underline{L}(\omega)\ \underline{P}\ \underline{L}^H(\omega)\} \qquad (4.6\text{-}1)$$

$\underline{L}(\cdot)$ is the matrix representation of $\mathfrak{L}(\cdot)$, and $\underline{P} = \underline{s}\ \underline{s}^H$, where $\underline{s}$ is the

vector of Fourier series expansion coefficients of the probing signal. If the Karhunen-Loève representation of $\underline{y}(\cdot,\cdot)$ is used, then

$$\underline{R}_y = \text{diag}[\lambda_1^{(y)} \ \lambda_2^{(y)} \ \ldots \ ] \tag{4.6-2}$$

which are the unavailable parameters in the estimator structure. The ith element in $\underline{R}_y$ is:

$$E\{|y_i(\omega)|^2\} = \lambda_i^{(y)} = \sum_m \sum_n E\{\ell_{im}(\omega) \ \ell_{in}^*(\omega)\} \ s(m) \ s^*(n) \tag{4.6-3}$$

This shows how the $\{\lambda_k^{(y)}\}$ terms are related to the second order statistics of the model parameters. Further insight is gained by evaluating the cross-correlations:

$$E\{\ell_{im}(\omega) \ \ell_{in}^*(\omega)\} = E\{< \mathfrak{L}\phi_m, \ \phi_i > < \mathfrak{L}\phi_n, \ \phi_i >^*\} \tag{4.6-4}$$

In terms of a random Green's function representation for $\mathfrak{L}(\cdot)$,

$$E\{< \mathfrak{L}\phi_m, \ \phi_i > < \mathfrak{L}\phi_n, \ \phi_i >^*\}$$

$$= E\{[\iint h(t, \tau, \omega) \ \phi_m(\tau) \ \phi_i^*(t) \ dt \ d\tau][\iint h(u, v, \omega) \ \phi_n(v) \ \phi_i^*(u) \ du \ dv]^*\}$$

$$= \iiiint E\{h(t, \tau, \omega) \ h^*(u, v, \omega)\} \ \phi_i^*(t) \ \phi_i(u) \ \phi_m(\tau) \ \phi_n^*(v) \ dt \ du \ d\tau \ dv$$

$$= \iiiint G(t, \tau, u, v) \ \phi_i^*(t) \ \phi_i(u) \ \phi_m(\tau) \ \phi_n^*(v) \ dt \ du \ d\tau \ dv \tag{4.6-5}$$

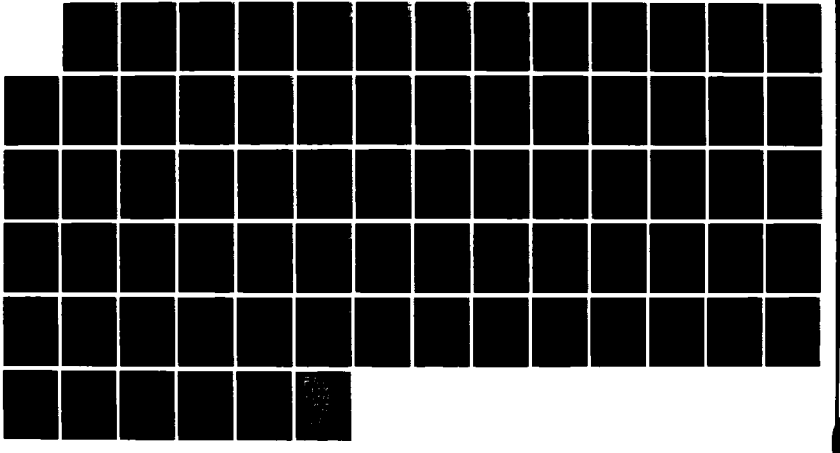
The expectation of the product $h(t, \tau, \omega)\, h^*(u, v, \omega)$ appears quite often in the stochastic system theory literature, and is important enough to warrant a name. $G(\cdot,\cdot,\cdot,\cdot)$ is called the stochastic Green's function of kernel $h(\cdot,\cdot,\cdot)$ [30]. It is a deterministic integral kernel which relates the statistical measures of the system output to those of its input. In this context, $G(\cdot,\cdot,\cdot,\cdot)$ relates the second order statistics of $\underline{y}(\cdot,\cdot)$ to the properties of probing signal $s(\cdot)$.

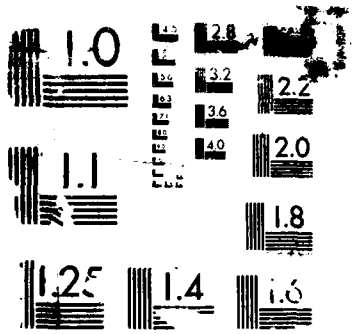
The four-fold integral (4.6-5) is the tensor representation of $G(\cdot,\cdot,\cdot,\cdot,)$ with respect to $\{\phi_k(\cdot)\}$. Therefore, Equation (4.6-5) can be interpreted as a parameterization of $G(\cdot,\cdot,\cdot,\cdot)$ with respect to the Karhunen-Loève basis $\{\phi_k(\cdot)\}$.

This result is significant for several reasons. First, it nails down the meaning of the somewhat nebulous expression, "Identification of the channel operator $\mathfrak{L}(\cdot)$." The cross-correlations of the matrix representation of $\mathfrak{L}(\cdot)$ can be interpreted as a parameterization of the very important stochastic Green's function $G(\cdot,\cdot,\cdot,\cdot)$. Estimating the numerical values of these parameters represents a systematic approach to stochastic operator identification. Furthermore, this approach allows a priori knowledge of the statistical properties of $\mathfrak{L}(\cdot)$ to be incorporated into the estimator-correlator structure. For example, it may be known that $\mathfrak{L}(\cdot)$ belongs to a certain class of channels characterized by certain forms of $G(\cdot,\cdot,\cdot,\cdot)$ or their mathematical equivalents, scattering function representations [5]. More often than not, these functions are parameterized by one or more variables which describe the particular channel under examination. It may be

AD-A189 130 ESTIMATION-CORRELATION MODELING AND IDENTIFICATION IN ADAPTIVE ARRAY PROC..(U) PENNSYLVANIA STATE UNIV STATE COLLEGE APPLIED RESEARCH LAB J A TAGUE ET AL. OCT 87 2/2

UNCLASSIFIED ARL/PSU/TR-87-005 N00024-85-C-6041 F/G 17/11 NL





can be determined a priori, or if it is possible to bound their values based on environmental constraints, this knowledge can be incorporated into the detector through Equation (4.6-5). Unfortunately, this interesting problem is beyond the scope of this study, and it is left as an issue which deserves further research.

4.7 Conclusions

The solutions for $\underline{Q}(\cdot,\cdot)$ and $\underline{G}(\cdot,\cdot)$ were derived and studied in detail. Spectral representations for the covariance and matrix filter kernels were used to obtain series solutions for the processor equations. We pointed out the relationship between these representations and the Karhunen-Loève expansion. In addition, it was seen that this approach sheds light onto several relevant theoretical issues and suggests how to implement the optimal processor.

The second order statistical properties of $\underline{f}(\cdot)$ are related to the $\{\lambda_k^{(y)}\}$ parameters. In principle, the channel identification process could be developed based on this relationship. However, this method has a fundamental limitation. The purpose of this study is to develop a channel identification scheme which can be incorporated into a practical array processor. Simplicity is of the essence in these applications, and clearly, the relationship between $\underline{R}_y(\cdot,\cdot)$ and $\underline{f}(\cdot)$ presented in the last section is much too complicated even if $\underline{R}_y$ is diagonal. A more compact representation for $\underline{f}(\cdot)$ is required. This problem will be examined in the next chapter.

Chapter 5

A CANONICAL EXPANSION APPROACH TO STOCHASTIC SYSTEM MODELING AND IDENTIFICATION

5.1 Introduction

This chapter examines the modeling and identification problem in further detail. In particular, we seek a more compact representation for $f(\cdot)$ in order to simplify Equation (4.6-3). Clearly, simplifying this expression is essential if the estimator-correlator is to be implemented. Of course, the requirements which were spelled out at the beginning of Chapter 4 must still be met. The representation for $f(\cdot)$ must be suitable for digital signal processing applications. Also, the channel identification should be signal independent. Once $f(\cdot)$ is identified using a probing signal $s(\cdot)$, we should be able to estimate the statistical measures of $\underline{y}(\cdot,\cdot)$ for any arbitrary transmitted signal. Finally, the representation should allow a priori knowledge of $f(\cdot)$ to be incorporated into the detector structure.

The derivation is presented in the next section. The solution immediately leads to a convenient series representation for $f(\cdot)$ developed in Section 5.3. These results are new and represent an original contribution to stochastic system theory. In addition, they provide interesting insights into several classical system identification theory issues. Using this representation, we obtain a very simple expression equivalent to Equation (4.6-3). It establishes an interesting connection among detection theory, estimation theory, and stochastic system identification theory which is examined in Section 5.5.

5.2 Derivation

In order for the results to be as general as possible, let us pose the modeling and identification problem in terms of a random Green's function representation for $\mathfrak{L}(\cdot)$, an input process $x(\cdot,\cdot)$, and an output process $y(\cdot,\cdot)$:

$$y(t, \omega) = \mathfrak{L}x(t, \omega) = \int_T h(t, \tau, \omega)\, x(\tau, \omega)\, d\tau \qquad (5.2\text{-}1)$$

$x(\cdot,\cdot)$ is a stochastic or deterministic probing signal whose properties are known and under our control. The final goal is to identify statistical measures of $h(\cdot,\cdot,\cdot)$ in terms of the known properties of $x(\cdot,\cdot)$ and the measurable properties of $y(\cdot,\cdot)$. We shall assume zero initial conditions, which is reasonable in the context of stochastic transmission channel identification.

Canonical expansions of the input and output covariance kernels will be used to identify the stochastic system. Let $\{\phi_k(\cdot)\}$ be an orthonormal basis simultaneously solving the following expressions:

$$\lambda_k^{(y)}\, \phi_k(t) = \int_T R_y(t, t_1)\, \phi_k(t_1)\, dt_1 \qquad (5.2\text{-}2)$$

$$\lambda_k^{(x)}\, \phi_k(t) = \int_T R_x(t, t_2)\, \phi_k(t_2)\, dt_2 \qquad (5.2\text{-}3)$$

In general, it is difficult to find simultaneous solutions to these eigenvalue equations. One approach is based on a generalized eigenvalue decomposition of $R_y(\cdot,\cdot)$ and $R_x(\cdot,\cdot)$ which will be described in detail later in the dissertation. There are numerically robust algorithms

available for solving the factorization problem. Furthermore, in a number of meaningful special cases, simultaneous solutions to Equations (5.2-2) and (5.2-3) exist. For example, if the input process $x(\cdot,\cdot)$ is wide sense stationary, and the random Green's function $h(\cdot,\cdot,\cdot)$ is time-invariant, then complex exponentials satisfy the equations. In addition, if either the input or output process can be approximated by a white noise process, then the problem reduces to solving only one of the eigenvalue equations. This approximation is reasonable since the probing signal $s(\cdot)$ is under our control. Good probing signal design can be exploited to simplify the representation and identification of $f(\cdot)$.

Next, recall that the system output $y(\cdot,\cdot)$ is given by

$$y(t, \omega) = \int_T h(t, \tau, \omega)\, x(\tau, \omega)\, d\tau$$

We shall make the physically realistic assumption that $h(\cdot,\cdot,\cdot)$ and $x(\cdot,\cdot)$ are uncorrelated. In terms of Equation (5.2-1), the system output covariance kernel $R_y(\cdot,\cdot)$ is:

$$R_y(t_1, t_2) = E\{[\int_T h(t_1, \tau, \omega)\, x(t, \omega)\, d\tau][\int_T h(t_2, s, \omega)\, x(s, \omega)\, ds]^*\}$$

Writing $R_y(\cdot,\cdot)$ in terms of expectations over the individual ensembles of $h(\cdot,\cdot,\cdot)$ and $x(\cdot,\cdot)$ gives:

$$R_y(t_1, t_2) = E_h\{[\int_T \int_T h(t_1,\tau,\omega)h^*(t_2,s,\omega)\, E_x\{x(t_1, \omega)x^*(t_2, \omega)\}\, d\tau\, ds\}$$

But

$$E\{x(\tau, \omega)\, x^*(s, \omega)\} = R_x(\tau, s) = \sum_{k=1}^{\infty} \lambda_k^{(x)}\, \phi_k(\tau)\, \phi_k^*(s)$$

and, of course,

$$R_y(t_1, t_2) = \sum_{k=1}^{\infty} \lambda_k^{(y)}\, \phi_k(t_1)\, \phi_k^*(t_2)$$

Therefore,

$$R_y(t_1, t_2) = \sum_{k=1}^{\infty} \lambda_k^{(y)}\, \phi_k(t_1)\, \phi_k^*(t_2)$$

$$= E_h\Big\{ \sum_{k=1}^{\infty} \lambda_k^{(x)} \int_T h(t_1, \tau, \omega)\, \phi_k(\tau)\, d\tau \int_T h^*(t_2, s, \omega)\, \phi_k^*(s)\, ds \Big\} \qquad (5.2\text{-}4)$$

Equating the kth terms in the expansion yields:

$$\int_T h(t_1, \tau, \omega)\, \phi_k(\tau)\, d\tau = h_k(\omega)\, \phi_k(t_1) \qquad (5.2\text{-}5)$$

$$\int_T h^*(t_2, s, \omega)\, \phi_k^*(s)\, ds = h_k^*(\omega)\, \phi_k^*(t_2)$$

Notice that the two preceding relationships are eigenvalue equations for $h(\cdot,\cdot,\cdot)$, which implies that in essence, a Karhunen-Loève expansion of $h(\cdot,\cdot,\cdot)$ is being performed. Substituting these relationships into Equation (5.2-4) and equating the k-th expansion coefficients gives:

$$R_y(t_1, t_2) = \sum_{k=1}^{\infty} \lambda_k^{(y)} \phi_k(t_1) \phi_k^*(t_2)$$

$$= \sum_{k=1}^{\infty} \lambda_k^{(x)} E_h \{h_k(\omega) h_k^*(\omega)\} \phi_k(t_1) \phi_k^*(t_2)$$

$$= \sum_{k=1}^{\infty} \lambda_k^{(x)} E\{|h_k(\omega)|^2\} \phi_k(t_1) \phi_k^*(t_2) \qquad (5.2\text{-}6)$$

This solves the identification problem. The eigenvalues $\{\lambda_k^{(y)}\}$ are known or can be estimated from the output measurements $y(\cdot,\cdot)$. Furthermore, the eigenvalues $\{\lambda_k^{(x)}\}$ are known because the process $x(\cdot,\cdot)$ is under our control. Therefore:

$$E\{|h_k(\omega)|^2\} = \lambda_k^{(h)} = \lambda_k^{(y)}/\lambda_k^{(x)} \qquad (5.2\text{-}7)$$

for $k = 1,2,\ldots$ This procedure identifies the second order statistics of the random Green's function $h(\cdot,\cdot,\cdot)$.

5.3 Representation of the Random Green's Function

Let us show how a convenient representation of $h(\cdot,\cdot,\cdot)$ can be found in terms of the basis $\{\phi_k(\cdot)\}$ and the coefficients $\{h_k(\cdot)\}$ described in the previous section.

Any bounded function defined on the product space $T \times T \times \Omega$ can be expanded as follows [22]:

$$h(t, \tau, \omega) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij}(\omega) \phi_i(t) \phi_j^*(\tau) \qquad (5.3\text{-}1)$$

where

$$h_{ij}(\omega) = \int_T \int_T h(t, \tau, \omega)\, \phi_i^*(t)\, \phi_j(\tau)\, dt\, d\tau \qquad (5.3\text{-}2)$$

and the functions

$$\Phi_{ij}(t, \tau) = \phi_i(t)\, \phi_j^*(\tau)$$

form an orthonormal basis for the product space T x T. The coefficients $\{h_{ij}(\cdot)\}$ are the matrix representation of the random Green's function $h(\cdot,\cdot,\cdot)$, or from a more abstract point of view, the matrix representation of stochastic operator $\mathfrak{L}(\cdot)$.

From Equation (5.2-5):

$$\int_T h(t, \tau, \omega)\, \phi_j(\tau)\, d\tau = h_j(\omega)\, \phi_j(t)$$

which implies that

$$h_{ij}(\omega) = \int_T h_j(\omega)\, \phi_j(t)\, \phi_i^*(t)\, dt = h_j(\omega)\, \delta_{ij} \qquad (5.3\text{-}3)$$

This result shows that the matrix representation of $\mathfrak{L}(\cdot)$ is diagonalized provided the covariance functions of $x(\cdot,\cdot)$ and $y(\cdot,\cdot)$ are simultaneously diagonalizable. Therefore, the random Green's function $h(\cdot,\cdot,\cdot)$ can be represented by a single sum series:

$$h(t, \tau, \omega) = \sum_{k=1}^{\infty} h_k(\omega)\, \phi_k(t)\, \phi_k^*(\tau) \qquad (5.3\text{-}4)$$

Estimating the means and variances of parameters $\{h_k(\cdot)\}$ is equivalent to identifying the mean and covariance functions of $h(\cdot,\cdot,\cdot)$.

5.4 Generalization to the Vector Case

The previous results can be generalized to vector processes. Suppose $\underline{x}(\cdot,\cdot)$ and $\underline{y}(\cdot,\cdot)$ are N x 1 stochastic processes. The covariance kernel of $\underline{y}(\cdot,\cdot)$ is an N x N matrix:

$$\underline{R}_y(t, u) = \sum_k \lambda_k^{(y)} \underline{\phi}_k(t) \underline{\phi}_k^H(u) \tag{5.4-1}$$

where the eigenvalues $\{\lambda_k^{(y)}\}$ and eigenfunctions $\{\underline{\phi}_k(\cdot)\}$ are solutions to the equation:

$$\lambda_k^{(y)} \underline{\phi}_k(y) = \int_T \underline{R}_y(t, u) \underline{\phi}_k(u)\, du \tag{5.4-2}$$

for k = 1,2,... Assuming simultaneous diagonalization of $\underline{R}_x(\cdot,\cdot)$ and $\underline{R}_y(\cdot,\cdot)$, the eigenvalues $\{\lambda_k^{(x)}\}$ and $\{\lambda_k^{(y)}\}$ are related as before:

$$\lambda_k^{(y)} = E\{|h_k(\omega)|^2\}/\lambda_k^{(x)} \tag{5.4-3}$$

for all k, and the matrix random Green's function $\underline{H}(\cdot,\cdot,\cdot)$ has the representation:

$$\underline{H}(t, \tau, \omega) = \sum_{k=1}^{\infty} h_k(\omega) \underline{\phi}_k(t) \underline{\phi}_k^H(t) \tag{5.4-4}$$

The second scheme to be considered is based on the representations for $\underline{x}(\cdot,\cdot)$ and $\underline{y}(\cdot,\cdot)$ derived in Section 3.4, where

$$\underline{x}(t, \omega) = \sum_{i=1}^{\infty} \sum_{j=1}^{N} x_{ij}(\omega)\, \phi_i(t)\, \underline{e}_j \tag{5.4-5}$$

$$\underline{y}(t, \omega) = \sum_{i=1}^{\infty} \sum_{j=1}^{N} y_{ij}(\omega)\, \phi_i(t)\, \underline{e}_j \tag{5.4-6}$$

Assuming the $\{y_{ij}(\cdot)\}$ coefficients are uncorrelated, then

$$\underline{R}_y(t, \omega) = \sum_{i=1}^{\infty} \sum_{j=1}^{N} \lambda_{ij}^{(y)}\, \phi_i(t)\, \phi_i^*(u)\, \underline{e}_j\, \underline{e}_j^H \tag{5.4-7}$$

But for a fixed i,

$$\sum_{j=1}^{N} \lambda_{ij}^{(y)}\, \underline{e}_j\, \underline{e}_j^H = \underline{U}(i)\, \underline{\Lambda}_i^{(y)}\, \underline{U}^H(i) \tag{5.4-8}$$

where the sum is written in matrix notation:

$$\underline{U}(i) = [\underline{e}_1 \ \ldots \ \underline{e}_N] \tag{5.4-9}$$

$$\underline{\Lambda}_i^{(y)} = \text{diag}[\lambda_{i1}^{(y)} \ \ldots \ \lambda_{iN}^{(y)}] \tag{5.4-10}$$

Substituting (5.4-8) into (5.4-7) leads to the following simplification

$$\underline{R}_y(t, u) = \sum_{i=1}^{\infty} \phi_i(t)\, \underline{U}(i)\, \underline{\Lambda}_i^{(y)}\, \underline{U}^H(i)\, \phi_i^*(u) \tag{5.4-11}$$

which will be used in the subsequent derivation.

For vector processes:

$$\underline{y}(t, \omega) = \int_T \underline{H}(t, \tau, \omega)\, \underline{x}(\tau, \omega)\, d\tau \qquad (5.4\text{-}12)$$

and

$$\underline{R}_y(t, u) = E_H\{\iint_{TT} \underline{H}(t, \tau, \omega)\, \underline{R}_x(\tau, v)\, \underline{H}^H(u, v, \omega)\, d\tau\, dv\} \qquad (5.4\text{-}13)$$

where an argument analogous to that in Section 5.3 is used. Again, it is assumed that $\underline{R}_x(\cdot,\cdot)$ and $\underline{R}_y(\cdot,\cdot)$ can be expressed in the representation (5.4-11) with respect to the same basis $\{\underline{\Phi}_{ij}(\cdot)\}$. Then

$$\underline{R}_x(\tau, v) = \sum_{i=1}^{\infty} \sum_{j=1}^{N} \lambda_{ij}^{(x)}\, \phi_i(\tau)\, \phi_i^*(v)\, \underline{e}_j\, \underline{e}_j^H$$

$$= \sum_{i=1}^{\infty} \phi_i(\tau)\, \underline{U}(i)\, \underline{\Lambda}_i^{(x)}\, \underline{U}^H(i)\, \phi_i^*(v) \qquad (5.4\text{-}14)$$

Substituting (5.4-11) into (5.4-15) and rearranging leads to the following:

$$\underline{R}_y(t,u) = E_H\{\sum_{i=1}^{\infty} [\int_T \underline{H}(t,\tau,\omega)\underline{U}(i)\phi_i(\tau)\, d\tau]\underline{\Lambda}_i^{(x)}[\int_T \underline{H}(u,v,\omega)\underline{U}(i)\phi_i(v)dv]^H\} \qquad (5.4\text{-}15)$$

Equating the k^{th} terms in (5.4-13) and (5.4-14) implies equality only if

$$\int_T \underline{H}(t, \tau, \omega)\, \underline{U}(k)\, \phi_k(\tau)\, d\tau = \phi_k(t)\, \underline{U}(k)\, \mathrm{diag}[h_{1k}(\omega) \ldots h_{Nk}(\omega)]$$

Then

$$\phi_k(t)\ \underline{U}(k)\ \underline{\Lambda}_k^{(y)}\ \underline{U}^H(k)\ \phi_k^*(u) = \phi_k(t)\ \underline{U}(k)\ \underline{\Lambda}_k^{(h)}\ \underline{\Lambda}_k^{(x)}\underline{U}^H(k)\ \phi_k^*(u)$$

for all k, and

$$\underline{\Lambda}_k^{(h)} = \underline{\Lambda}_k^{(y)}\ \underline{\Lambda}_k^{(x)-1} \tag{5.4-16}$$

Each matrix in Equation (5.4-16) is diagonal; therefore, the identification is easy to perform. Moreover, the representation of $\underline{H}(\cdot,\cdot,\cdot)$ with respect to $\{\underline{\phi}_{ij}(\cdot)\}$ reduces to:

$$\underline{H}(t,\ \tau,\ \omega) = \sum_{i=1}^{\infty}\ \sum_{j=1}^{N} h_{ij}(\omega)\ \phi_i(t)\ \phi_i^*(\tau)\ \underline{e}_j\ \underline{e}_j^H \tag{5.4-17}$$

which is the matrix equivalent of (5.3-4).

5.5 Discussion

The results from Sections 5.2 through 5.4 show how Karhunen-Loève expansions can be applied to the system modeling and identification problem. Expanding the random Green's function in terms of the Karhunen-Loève basis is a fundamental approach to stochastic system modeling. It simplifies the identification problem, and gives a representation for $f(\cdot)$ which can easily be incorporated into the estimator-correlator structure. One can interpret the results as a transformation of the matrix parameters from an arbitrary basis into the Karhunen-Loève basis, where the identification is easier to perform.

The results provide several interesting new insights into the stochastic system identification problem. Applying Equation (5.2-7) to the original problem of simplifying Equation (4.6-3) yields

$$\lambda_k^{(y)} = E\{|\ell_k(\omega)|^2\}\ \lambda_k^{(s)} \tag{5.5-1}$$

where $\{\lambda_k^{(s)}\}$ are the eigenvalues of $\underline{P}$. This result is deceptively simple looking; actually, it ties together ideas from several disciplines. The representation for $f(\cdot)$ was derived in order to simplify the relationship between the $\{\lambda_k^{(y)}\}$ terms and the $\{\lambda_k^{(s)}\}$ terms, in other words, to simplify the channel identification problem. The $\{\lambda_k^{(y)}\}$ terms are needed to calculate the minimum mean-square estimate of the channel output used to compute the likelihood ratio. Of course, the key to obtaining (5.5-1) is simultaneous diagonalization of $\underline{P}$ and $\underline{R}_y$, which in this context depends on proper design of the probing signal $s(\cdot)$, or in the case of more general signals, the use of numerically robust algorithms solving a generalized eigenvalue problem. The representation for $f(\cdot)$ meets the requirements defined at the beginning of the chapter. It is well-suited for digital processing of the array data, and it allows a priori knowledge of the scattering channel to be incorporated into the estimator-correlator structure.

Equation (5.5-1) also gives new insight into the meaning of the phrase, "For system identification, a probing signal must be sufficiently rich [37] [38]." In terms of Equation (5.5-1), it means that for each eigenvalue $\lambda_k^{(y)}$, the probing signal eigenvalue $\lambda_k^{(s)}$ must be large enough so that numerical errors do not occur while performing the

division needed to estimate the $E\{|h_k(\omega)|^2\}$ terms. This is a key to good probing signal design, and furthermore, it suggests posing the signal design problem as an inverse eigenvalue problem [35].

5.6 Conclusions

The Karhunen-Loève expansion represents a fundamental approach to stochastic system modeling and identification. Simultaneous diagonalization of the input and output process covariance kernels can be accomplished through proper probing signal design or generalized eigenvalue decomposition. The representation for $\mathfrak{L}(\cdot)$ simplifies the identification problem and gives insight into classical identification theory issues.

Chapter 6

NUMERICAL ISSUES

6.1 Introduction

At this point in the dissertation, the mathematical derivations needed to solve for the processor kernels have been completed. We have asserted that orthonormal representations of the covariance kernels, the processor kernels, and stochastic operator $\mathfrak{L}(\cdot)$ can be used to obtain the solutions. However, several outstanding issues still must be resolved before the optimal structure can actually be implemented. In particular, we have yet to show how the Karhunen-Loève eigenfunctions and eigenvalues can be calculated from array measurements. This problem must be examined in detail if we are to go beyond formal manipulations of infinite series to a working system. In conjunction with this issue, the numerical difficulties inherent in *any adaptive signal processing* system must be overcome. How do numerical difficulties arise?

The answer to this question relates to the nature of adaptive systems. Recall that an adaptive system is a learning or self-optimizing machine which adjusts its response according to the statistical properties of its surroundings [40]. This is where numerical difficulties can occur. For example, second order statistical information is usually estimated by post-multiplying a data matrix by its Hermitian transpose, which causes loss of numerical precision when the arithmetic operations are carried out on finite-precision hardware. The

ability to implement an equivalent adaptive processor in hardware without performing the squaring operation is very desirable.

Since a priori knowledge of the covariance kernels is incomplete, they must be estimated from array data. Therefore, the numerical problems as described above must be taken into account while designing the processor.

The purpose of this chapter is to solve the computational and adaptive implementation problems required to construct the optimal processor. We begin by returning to the inverse noise covariance kernel $\underline{Q}(\cdot,\cdot)$, and suggest an approach for computing the Karhunen-Loève basis. The solution requires calculating matrix products, an operation which should be avoided whenever possible. The singular value decomposition can be used to solve an equivalent estimation problem, bypassing the squaring step altogether.

Next, we shall turn our attention to the estimator branch $\underline{G}(\cdot,\cdot)$. The solution is based on simultaneous diagonalization of $\underline{R}_y(\cdot,\cdot)$ and $\underline{R}_1(\cdot,\cdot)$ using generalized eigenvalue decomposition. Once more, a matrix squaring operation appears in the solution formulation. Can an equivalent processing system be realized without squaring?

The answer to this important question is yes, and the processing is based on a matrix decomposition which is just now appearing in the numerical signal processing literature [41]. It is called the CS (cosine-sine) decomposition, and it provides the means to compute $\underline{G}(\cdot,\cdot)$ directly from array data, making the matrix multiplication unnecessary.

Furthermore, it will be proved that this solution is numerically equivalent to decorrelating the expansion coefficients of $\underline{n}(\cdot,\cdot)$ followed by a Karhunen-Loève transformation.

Finally, we will show how to obtain the data which are needed to estimate $\underline{R}_N(\cdot,\cdot)$, $\underline{R}_1(\cdot,\cdot)$, and $\underline{R}_y(\cdot,\cdot)$. It will be seen that $\underline{R}_y(\cdot,\cdot)$ must be obtained at a high signal-to-noise ratio, as this step represents system identification.

6.2 The Inverse Noise Covariance Kernel

6.2.1 Example

The computational issues involved with implementation can be illustrated by considering the following example. Suppose that $\underline{n}(\cdot,\cdot)$ is a wide-sense stationary, periodic stochastic process. Find the Karhunen-Loève basis and $\underline{Q}(\cdot,\cdot)$.

We begin by selecting a basis in the form derived in Section 3.4:

$$\underline{\Phi}_{ij}(t) = \phi_i(t)\, \underline{u}_j \tag{6.2-1}$$

for $j = 1,2\ldots, M$, where M is the number of sensors, and for all i. $\underline{n}(\cdot,\cdot)$ can be expanded as follows:

$$\underline{n}(t,\omega) = \sum_{i=-\infty}^{\infty} \sum_{j=1}^{N} n_{ij}(\omega)\, \phi_i(t)\, \underline{u}_j$$

and $\underline{Q}(\cdot,\cdot)$ has the representation:

$$\underline{Q}(t,u) = \sum_{i=-\infty}^{\infty} \sum_{j=1}^{N} \frac{1}{\lambda_{ij}^{(n)}}\, \phi_i(t)\, \phi_i^*(u)\, \underline{u}_j\, \underline{u}_j^H \tag{6.2-2}$$

The basis elements must be orthogonal:

$$\langle \underline{\Phi}_{ij}, \underline{\Phi}_{k\ell} \rangle = \delta_{ik}\, \delta_{j\ell}$$

and the expansion coefficients must be uncorrelated:

$$E\{n_{ij}(\omega)\, n^{*}_{k\ell}(\omega)\} = \lambda^{(n)}_{ij}\, \delta_{ik}\, \delta_{j\ell}$$

These identities suggest that we attempt to solve the integral equation

$$\lambda_{mn}\, \underline{\Phi}_{mn}(t) = \int_{T} \underline{R}_{N}(t,s)\, \underline{\Phi}_{mn}(s)\, ds \qquad (6.2\text{-}3)$$

for all m, and for n = 1,..., M. Substituting Equation (6.2-1) into (6.2-3) gives:

$$\lambda_{mn}\, \phi_{m}\, \underline{u}_{n} = \int_{T} \underline{R}_{N}(t,s)\, \phi_{m}(s)\, \underline{u}_{n}\, ds$$

Assuming $\underline{n}(\cdot,\cdot)$ is periodic with $T = 2\pi$, the integral equation becomes

$$\lambda_{mn}\, \phi_{n}(t)\, \underline{u}_{n} = \int_{-\pi}^{\pi} \underline{R}_{N}(\tau)\, \phi_{m}(t-\tau)\, \underline{u}_{n}\, d\tau \qquad (6.2\text{-}4)$$

The proper choice for the scalar basis is the complex exponentials:

$$\phi_{m}(t) = \frac{1}{\sqrt{2\pi}} \exp\{jmt\}$$

For a given m, Equation (6.2-4) becomes:

$$\lambda_{mn}\underline{u}_n = \int_{-\infty}^{\infty} \underline{R}_N(\tau)\, e^{-jm\tau}\, \underline{u}_n\, d\tau = \underline{S}_N(m\omega_0)\, \underline{u}_n$$

where $\underline{S}_N(\cdot)$ is the power spectral density matrix of $\underline{n}(\cdot,\cdot)$, and ω_0 is equal to one. Therefore, the equation

$$\lambda_{mn}\underline{u}_n = \underline{S}_N(m)\underline{u}_n \tag{6.2-5}$$

must be solved for eigenvalues $\{\lambda_{mn}\}$ and eigenvectors $\{\underline{u}_n\}$ to obtain the Karhunen-Loève basis. Actually, since the eigenvectors $\{\underline{u}_n\}$ are also a function of m, Equation (6.2-5) should be:

$$\lambda_{mn}\, \underline{u}_{mn} = \underline{S}_N(m)\, \underline{u}_{mn} \tag{6.2-6}$$

Solving Equation (6.2-6) provides the Karhunen-Loève basis. Of course, the wide-sense stationary, periodic signal assumptions were only made for purposes of illustration. The preceding results can be generalized to nonstationary processes as well.

6.2.2 A Possible Processing Scheme

The results from Section 6.2.1 suggest a processing scheme illustrated in Figure 6-1. Since the scalar basis is the set of complex exponentials, the first step is to compute the discrete Fourier transform of the array data given noise alone. Next, the power spectral density matrix at each harmonic $m\omega_0$ is estimated by averaging over the array data. Finally, the vector basis functions are calculated by

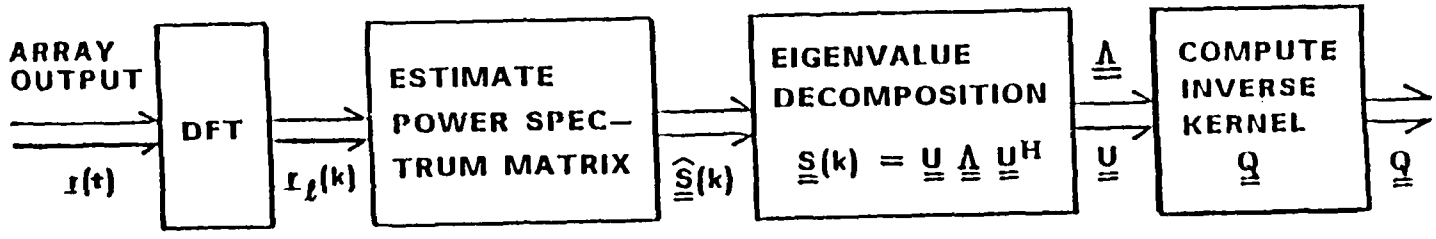


Figure 6-1. Calculating the Karhunen-Loéve Basis

performing an eigenvalue decompcsition of each estimated power spectral density matrix.

While this sequence of steps is mathematically correct, it is not advisable for the reasons given in Section 6.1 [42]. The power spectral density matrix estimation involves a matrix squaring operation prior to the eigenvalue decomposition step. Since this reduces the precision of the final answer, an equivalent procedure avoiding this step is needed. This is where the singular value decomposition shall be introduced.

6.2.3 Applying the Singular Value Decomposition

The loss of numerical precision can be avoided if the singular value decomposition (SVD) [39] is used instead of eigenvalue decomposition. The first step in an equivalent processing operation is to calculate the discrete Fourier transform as before. Next, consider a sequence of vectors $\{\underline{r}(k)\}$. Arrange them sequentially in the N x L matrix $\underline{A}(k)$:

$$\underline{A}(k) = [\underline{r}_1(k)\ \underline{r}_2(k)\ \ldots\ \underline{r}_L(k)] \qquad (6.2\text{-}7)$$

where $\underline{r}_i(k)$ is the i-th M x 1 vector, and L is the number of measurements. The notation is simplified by dropping the notation k. Then $\underline{A}$ has the following decomposition:

$$\underline{A} = \underline{U}\ \underline{\Sigma}\ \underline{V}^H \qquad (6.2\text{-}8)$$

where $\underline{U}$ is an M x M unitary matrix, $\underline{V}$ is an L x L unitary matrix, and $\underline{\Sigma}$ is a diagonal matrix whose elements are the singular values of $\underline{A}$. $\underline{A}$ can be written in terms of the singular values $\{\sigma_i\}$ and the columns of $\underline{U}$ and $\underline{V}$:

$$\underline{A} = \sum_{i=1}^{r} \sigma_i \, \underline{u}_i \, \underline{v}_i^H \qquad (6.2\text{-}9)$$

r is the number of non-zero singular values of $\underline{A}$, which is also the rank of $\underline{A}$. Furthermore, the following relationships between the columns of $\underline{U}$, columns of $\underline{V}$, and $\underline{A}$ are met:

$$\underline{A} \, \underline{v}_m = \sigma_m \, \underline{u}_m \qquad (6.2\text{-}10)$$

$$\underline{A}^H \, \underline{u}_m = \sigma_m \, \underline{v}_m \qquad (6.2\text{-}11)$$

An eigenvector-eigenvalue relationship is found by eliminating $\underline{v}_m$ from Equations (6.2-10) and (6.2-11):

$$\underline{A} \, \underline{A}^H \, \underline{u}_m = \sigma_m^2 \, \underline{u}_m \qquad (6.2\text{-}12)$$

$\underline{A} \, \underline{A}^H$ is a Hermitian matrix, and $\sigma_m^2 = \lambda_m$.

The point of this discussion is that the eigenvalue decomposition can be calculated by solving an equivalent singular value decomposition. The power spectral density matrix estimates are given in the form

$$\underline{A}(k) \, \underline{A}^H(k) \qquad (6.2\text{-}13)$$

at each frequency k. It is clear from (6.2-10) and (6.2-12) that the eigenvectors of this product are the same as the left singular vectors of $\underline{A}(k)$. Therefore, the eigenvectors and eigenvalues of (6.2-13) can be obtained directly from transformed array data by calculating a singular value decomposition. The procedure is illustrated in Figure 6-2.

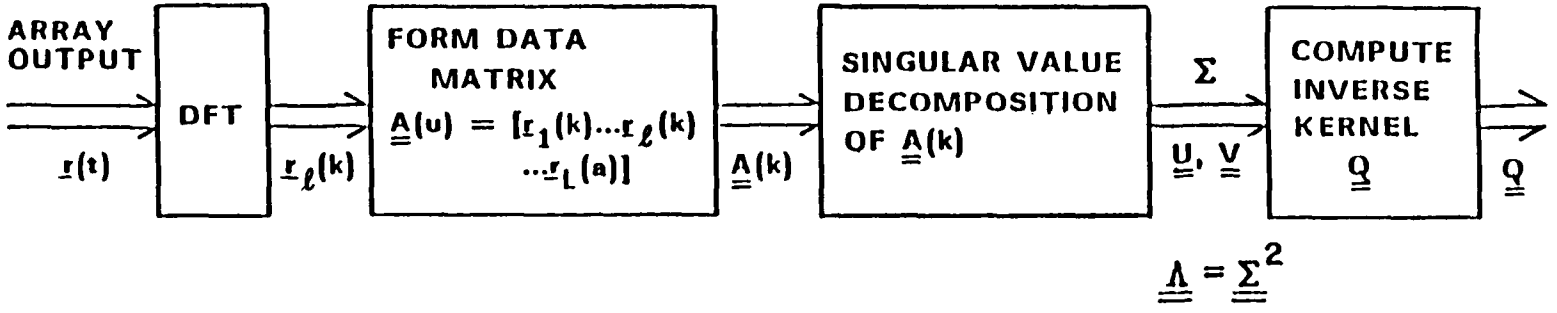


Figure 6-2. A Better Processing Scheme

The previous discussion suggests how the results in Section 2.5 can be implemented adaptively. Calculating the Karhunen-Loève basis and expansion coefficients from an arbitrary expansion requires an eigenvalue decomposition of the Fourier series coefficient covariance matrix. If this matrix is unknown, it can be estimated from array data:

$$\hat{\underline{R}}_x = \underline{A}^H \underline{A}$$

where $\underline{A}$ is a data matrix of Fourier coefficients computed over L measurements of the array output. Rather than computing the eigenvalue decomposition of $\hat{\underline{R}}_x$, a singular value decomposition of $\underline{A}$ is calculated. The left singular vector matrix $\underline{U}$ represents the linear transformation needed to calculate the Karhunen-Loève basis and series coefficients.

6.3 The Estimator Kernel

6.3.1 Introduction

Next, the numerical issues involved with solving the equation

$$\int_T \underline{R}_1(t,u)\, \underline{G}(u,z)\, du = \underline{R}_y(t,z) \qquad (6.3\text{-}1)$$

for $\underline{G}(\cdot,\cdot)$ will be examined. A numerically robust solution will be derived based on the CS matrix decomposition.

6.3.2 Basic Numerical Approach

In Section 4.4 we demonstrated how the solution for $\underline{G}(\cdot,\cdot)$ was constructed by data prewhitening followed by an eigenvalue decomposition of the modified covariance kernels $\underline{R}_1'(\cdot,\cdot)$ and $\underline{R}_y'(\cdot,\cdot)$. Here, the problem will be approached from a different perspective. What we seek is a linear transformation which diagonalizes $\underline{R}_1(\cdot,\cdot)$ and $\underline{R}_y(\cdot,\cdot)$ directly, skipping the prewhitening step entirely. In terms of matrix representations, this problem can be posed as follows. Find a linear transformation $\underline{X}$ such that $\underline{R}_1$ and $\underline{R}_y$ are diagonalized simultaneously

$$\underline{X}^H \underline{R}_1 \underline{X} = \text{diag}[\alpha_1 \ \alpha_2 \ \ldots \ ] \tag{6.3-2}$$

$$\underline{X}^H \underline{R}_y \underline{X} = \text{diag}[\beta_1 \ \beta_2 \ \ldots \ ] \tag{6.3-3}$$

Provided a non-singular $\underline{X}$ exists, then:

$$\underline{R}_1 = \underline{X}^{-H} \underline{\Lambda}_\alpha \underline{X}^{-1}$$

where the $-H$ notation denotes inversion followed by Hermitian transpose, and similarly,

$$\underline{R}_y = \underline{X}^{-H} \underline{\Lambda}_\beta \underline{X}^{-1}$$

Since $\underline{G}$ is given by:

$$\underline{G} = \underline{R}_y \underline{R}_1^{-1}$$

in terms of $\underline{X}$, $\underline{\Lambda}_\alpha$, and $\underline{\Lambda}_\beta$, $\underline{G}$ is written:

$$\underline{G} = \underline{R}_y \underline{R}_1^{-1} = \underline{X}^{-H} \underline{\Lambda}_\beta \underline{X}^{-1} \underline{X} \underline{\Lambda}_\alpha^{-1} \underline{X}^H = \underline{X}^{-H} \underline{\Lambda}_\beta \underline{\Lambda}_\alpha^{-1} \underline{X}^H$$

$$= \underline{X}^{-H} \text{ diag}[\beta_1/\alpha_1 \ \ldots\ ] \underline{X}^H \qquad (6.3\text{-}4)$$

The necessary and sufficient conditions for simultaneous diagonalization of two positive definite, Hermitian matrices are well understood [43]. It can be shown that there exists a non-singular, unique matrix $\underline{X}$ which diagonalizes $\underline{R}_1$ and $\underline{R}_y$. Furthermore, the computational aspects of this problem have been studied extensively by numerical analysts. In this field, the problem is called the *generalized eigenvalue problem of the matrix pencil* $\underline{R}_1 - \lambda \underline{R}_y$ [39]. The fact that the columns of $\underline{X}$ correspond to the generalized eigenvectors of $\underline{R}_1$ and $\underline{R}_y$ can be demonstrated by rewriting Equations (6.3-2) and (6.3-3):

$$\underline{R}_1 \underline{X} = \underline{X}^{-H} \text{ diag}[\alpha_1 \ \alpha_2 \ \ldots\ ]$$

$$\underline{R}_y \underline{X} = \underline{X}^{-H} \text{ diag}[\beta_1 \ \beta_2 \ \ldots\ ]$$

and by defining $\underline{E} = \underline{X}^{-H}$:

$$\underline{R}_1 \underline{X} = \underline{E} \text{ diag}[\alpha_1 \ \alpha_2 \ \ldots\ ]$$

$$\underline{R}_y \underline{X} = \underline{E} \text{ diag}[\beta_1 \ \beta_2 \ \ldots\ ]$$

For the kth columns of $\underline{X}$ and $\underline{E}$:

$$\underline{R}_1 \underline{x}_k = \alpha_k \underline{e}_k$$

$$\underline{R}_y \underline{x}_k = \beta_k \underline{e}_k$$

Rearranging these relations leads to the following:

$$\underline{R}_1 \, \underline{x}_k = (\alpha_k/\beta_k) \, \underline{R}_y \, \underline{x}_k = \lambda_k \, \underline{R}_y \, \underline{x}_k$$

which shows that the rows of $\underline{X}$ are the generalized eigenvectors of $\underline{R}_1$ and $\underline{R}_y$, and that the $\{\alpha_k/\beta_k\}$ terms are the generalized eigenvalues of $\underline{R}_1$ and $\underline{R}_y$.

The preceding discussion suggests that $\underline{G}(\cdot,\cdot)$ can be calculated directly by solving a generalized eigenvector problem, where $\underline{G}(\cdot,\cdot)$ is expressed directly in terms of a generalized spectral representation. But actually, this solution is equivalent to that presented in Section 4.4. This can be seen by defining a filter operator $\underline{C}$ as the matrix performing Gram-Schmidt orthogonalization of the noise process Fourier coefficients:

$$\underline{n}'(\omega) = \underline{C} \, \underline{n}(\omega) \tag{6.3-5}$$

where

$$E\{\underline{n}'(\omega) \, \underline{n}'^{H}(\omega)\} = \underline{I}$$

$\underline{C}$ can be written in upper triangular form, and given a positive definite, Hermitian noise covariance kernel $\underline{R}_N$, $\underline{C}$ is unique [44]. Now, when the array data $\underline{r}(\cdot)$ are prewhitened by $\underline{C}$:

$$\underline{r}'(\omega) = \underline{C} \, \underline{r}(\omega) = \underline{C} \, \underline{y}(\omega) + \underline{C} \, \underline{n}(\omega) = \underline{y}'(\omega) + \underline{n}'(\omega)$$

the covariance kernels of $\underline{r}'(\cdot)$ and $\underline{y}'(\cdot)$ are:

$$\underline{R}_1' = E\{\underline{r}'(\omega)\ \underline{r}'^H(\omega)\} = \underline{C}\ \underline{R}_1\ \underline{C}^H = \underline{C}\ \underline{R}_y\ \underline{C}^H + \underline{I} = \underline{R}_y' + \underline{I}$$

Both $\underline{R}_1'$ and $\underline{R}_y'$ are positive definite and Hermitian, their eigenvectors are the same, and they are unique:

$$\underline{R}_1' = \underline{U}\ \underline{\Lambda}_1\ \underline{U}^H \tag{6.3-6}$$

$$\underline{R}_y' = \underline{U}\ \underline{\Lambda}_y\ \underline{U}^H \tag{6.3-7}$$

Rearranging Equations (6.3-6) and (6.3-7) yields:

$$\underline{U}^H\ \underline{R}_1'\ \underline{U} = \text{diag}[\lambda_1^{(r')}\ \ldots\ ]$$

$$\underline{U}^H\ \underline{R}_y'\ \underline{U} = \text{diag}[\lambda_1^{(y')}\ \ldots\ ]$$

But in terms of $\underline{R}_1$, $\underline{R}_y$, and $\underline{C}$:

$$\underline{U}^H\ \underline{C}\ \underline{R}_1\ \underline{C}^H\ \underline{U} = \text{diag}[\lambda_1^{(r')}\ \ldots\ ]$$

$$\underline{U}^H\ \underline{C}\ \underline{R}_y\ \underline{C}^H\ \underline{U} = \text{diag}[\lambda_1^{(y')}\ \ldots\ ]$$

The matrices $\underline{U}$ and $\underline{C}$ are unique, therefore:

$$\underline{X}^H = \underline{U}^H\ \underline{C} = \underline{Q}\ \underline{R} \tag{6.3-8}$$

which equates the eigenvectors of $\underline{R}_1'$ and $\underline{R}_y'$ with the orthonormal columns $\underline{Q}$ of the QR factorization of $\underline{X}^H$, and the whitening filter $\underline{C}$ with matrix $\underline{R}$. The $\{\alpha_k\}$ and $\{\beta_k\}$ terms are the eigenvalues of $\underline{R}_1'$ and $\underline{R}_y'$ respectively. Therefore, solving the generalized eigenvalue problem is equivalent to decorrelating $\underline{n}(\cdot)$ followed by a Karhunen-Loève transformation.

6.3.3 Implementing the Estimator Branch with the CS Decomposition

We have seen how to implement the estimator branch using the generalized eigenvalue decomposition, and in principle, this part of the problem is solved. In this section, we consider how implementation can be accomplished in a numerically robust fashion.

Maximum likelihood estimates for $\underline{R}_1$ and $\underline{R}_y$ can be constructed by averaging over independent array measurements. The procedure begins by calculating the appropriate Fourier series expansion coefficients:

$$\begin{aligned} r_1(\omega) &= \langle \underline{r}, \underline{\phi}_1 \rangle \\ &\vdots \\ r_k(\omega) &= \langle \underline{r}, \underline{\phi}_k \rangle \end{aligned}$$

and so on. This step is repeated over L measurements, and a data matrix $\underline{A}$ is constructed:

$$\underline{A}^H = [\underline{r}_1 \ \underline{r}_2 \ \cdots \ \underline{r}_k \ \cdots \ ]$$

The kth column of $\underline{A}^H$ contains the L coefficients calculated with respect to the kth basis function. Finally, the estimated covariance matrix $\hat{\underline{R}}_1$ is formed by calculating the product

$$\hat{\underline{R}}_1 = \underline{A}^H \underline{A} \tag{6.3-9}$$

The signal alone covariance kernel must also be estimated from array measurements. Of course, the signal $\underline{s}(\cdot)$ is always obscured by additive measurement noise, meaning that it is not possible to estimate $\underline{R}_y$ directly from data. However, if the signal-to-noise ratio is large enough, the array measurements can be made close to signal alone:

$$\underline{r}(\omega) = \underline{y}(\omega) + \underline{n}(\omega) \approx \underline{y}(\omega)$$

Since the probing signal is under our control, high signal-to-noise ratio measurements can be obtained by adjusting the probing signal energy. By repeating the processing steps described above, a data matrix $\underline{B}$ which approximates signal alone measurements is formed:

$$\underline{B}^H = [\underline{y}_1 \ \underline{y}_2 \ \cdots \ ]$$

and $\underline{R}_y$ is estimated by

$$\hat{\underline{R}}_y = \underline{B}^H \underline{B} \tag{6.3-10}$$

Finally, the generalized eigenvectors and eigenvalues can be calculated by solving the system

$$\underline{A}^H \underline{A} \ \underline{x}_k = \lambda_k \ \underline{B}^H \underline{B} \ \underline{x}_k \tag{6.3-11}$$

for its eigenvalues and eigenvectors.

In order to solve Equation (6.3-11) without squaring, we propose a new approach which is just now becoming known in the signal processing

community [41] [45]. The solution uses the CS decomposition, which will be described in the next section.

6.3.4 Solution

The CS (cosine-sine) matrix decomposition arises naturally in the context of the $\underline{A}^H \underline{A} - \lambda \underline{B}^H \underline{B}$ generalized eigenvalue problem. This system can be solved by CS decomposition directly in terms of $\underline{A}$ and $\underline{B}$. Suppose

$$\begin{bmatrix} \underline{A} \\ \underline{B} \end{bmatrix} = \begin{bmatrix} \underline{Q}_1 \\ \underline{Q}_2 \end{bmatrix} \underline{R} \tag{6.3-12}$$

is the QR decomposition of $\underline{A}$ over $\underline{B}$. Writing $\underline{Q}_1$ and $\underline{Q}_2$ in terms of their singular value decompositions gives

$$\underline{Q}_1 = \underline{U}_1 \, \underline{C} \, \underline{V} \tag{6.3-13}$$

$$\underline{Q}_2 = \underline{U}_2 \, \underline{S} \, \underline{V} \tag{6.3-14}$$

where

$$\underline{C} = \text{diag}(c_i)$$

and

$$\underline{S} = \text{diag}(s_i)$$

The $\{c_i\}$ and $\{s_i\}$ terms are non-negative, $\underline{U}_1$, $\underline{U}_2$, and $\underline{V}$ are unitary, and

$$\underline{C}^2 + \underline{S}^2 = \underline{I}$$

By setting

$$\underline{X} = (\underline{V}^H \underline{R})^{-1} \tag{6.3-15}$$

$$\text{diag}[\alpha_1 \ \ldots \ \alpha_N] = \underline{C}^T \underline{C} \tag{6.3-16}$$

$$\text{diag}[\beta_1 \ \ldots \ \beta_N] = \underline{S}^T \underline{S} \tag{6.3-17}$$

then

$$\underline{X}^H (\underline{A}^H \underline{A}) \underline{X} = \text{diag}[\alpha_1 \ \ldots \ \alpha_N]$$

$$\underline{X}^H (\underline{B}^H \underline{B}) \underline{X} = \text{diag}[\beta_1 \ \ldots \ \beta_N]$$

and it follows that $\underline{X}$ is the matrix simultaneously diagonalizing $\underline{A}^H \underline{A}$ and $\underline{B}^H \underline{B}$. Therefore, using this approach, the estimates for $\underline{X}$, the generalized eigenvalues, and $\underline{G}$ can be calculated directly from array data without performing squaring.

The preceding discussion leads to the following processing steps needed to calculate $\underline{G}$: Form data matrices $\underline{A}$ and $\underline{B}$ based on appropriate Fourier series representations, such as the discrete Fourier transform. Next, calculate the QR decomposition of the $\underline{A}$ over $\underline{B}$ matrix. Third, calculate the singular value decompositions of $\underline{Q}_1$ and $\underline{Q}_2$. Fourth, form $\underline{X}$ from the $\underline{V}$ and $\underline{R}$ matrices obtained in steps three and one respectively. Finally,

$$\underline{G} = \underline{X}^{-H}[\underline{S}^T\underline{S}][\underline{C}^T\underline{C}]^{-1} \underline{X}^H \tag{6.3-18}$$

6.4 Adaptive Covariance Kernel Estimation

Several details involved with covariance kernel estimation must be worked out. Since these estimates depend on array data, it is clear that the proper information must be used to construct the estimates. For example, $\underline{R}_N(\cdot,\cdot)$ must be constructed from measurements given no signal present. Similarly, both signal and noise should be present in the measurements used to generate data matrix $\underline{A}$. However, these steps present a dilemma, since the processor is designed to perform detection. How can the covariance kernel estimates be constructed without a priori knowledge of the correct hypothesis?

The answer to this question relates to the fundamental nature of adaptive processing systems. Building the covariance kernel estimates is a learning step in which a priori knowledge of the underlying data structure must be known. Therefore, in order to determine $\underline{R}_N(\cdot,\cdot)$ and $\underline{R}_1(\cdot,\cdot)$, we will assume that a priori knowledge of the correct hypothesis is available.

In the active detection problem, one can use the scheme in Figure 6-3 [4]. Here T' is the return travel time of the probing signal, and T is once again the observation interval. The noise covariance kernel estimate can be formulated during the interval T', and used in subsequent detection. The signal plus noise kernel is estimated over the time interval T.

Calculating the signal alone kernel is not quite as straightforward, since signal alone measurements are not available. However,

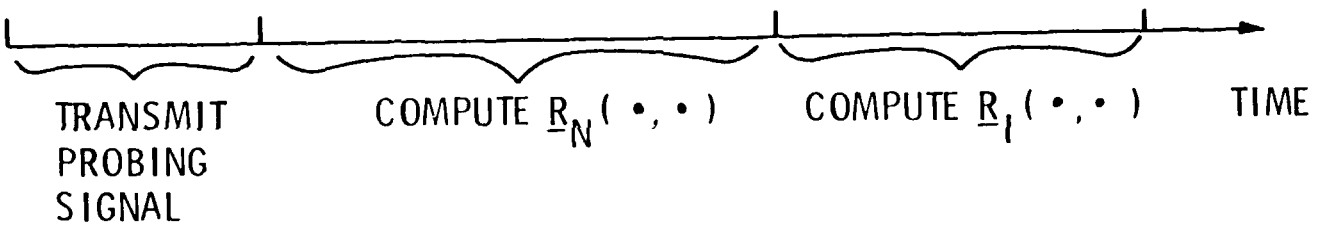


Figure 6-3. Estimation of the Covariance Kernels

the probing signal $s(\cdot)$ is under our control, and if its energy is large enough, we can argue that:

$$\underline{r}(t, \omega) = \underline{fs}(t) + \underline{n}(t, \omega) \approx \underline{fs}(t) \qquad (6.4\text{-}1)$$

and by proper scaling of the data measurements, the kernel $\underline{R}_y(\cdot,\cdot)$ corresponding to the data measured over the time interval T can be estimated. Clearly, a high signal-to-noise ratio is needed during this step. This is reasonable from a system identification point of view, since $\underline{R}_y(\cdot,\cdot)$ represents the second order statistics of an unknown linear system. System identification is fundamentally a high signal-to-noise ratio process. Of course, once the data matrix $\underline{B}$ is formed at a high signal-to-noise ratio, the solution can be incorporated into the detection process, which is generally carried out in low signal-to-noise situations.

6.5 Conclusions

We have examined the adaptive implementation issues in some detail and have proposed processing schemes based on robust, state-of-the-art numerical algorithms. Our results have established interesting connections between several canonical matrix decompositions and the spectral representations of $\underline{Q}(\cdot,\cdot)$ and $\underline{G}(\cdot,\cdot)$ presented in Chapter 4.

It is clear that the singular value decomposition is a very useful numerical processing tool. The eigenvectors and eigenvalues of the estimated power spectral density matrices were obtained directly from data matrices without performing a squaring operation. In another context, this decomposition could be used to estimate covariance matrix

eigenvectors from array measurements. The advantages of this approach are well documented [46].

We proposed using the CS decomposition to solve for $\underline{G}(\cdot,\cdot)$ numerically. This method is equivalent to noise coefficient decorrelation followed by a Karhunen-Loève transformation. Furthermore, the CS decomposition solves for the generalized eigenvectors and eigenvectors of $\underline{R}_I(\cdot,\cdot)$ and $\underline{R}_y(\cdot,\cdot)$ directly from data matrices, once again making matrix multiplication unnecessary.

Finally, we showed how to adaptively estimate the covariance kernels from array measurements. This part of the problem must be interpreted as a learning procedure performed prior to detection. $\underline{R}_y(\cdot,\cdot)$ must be obtained at a high signal-to-noise ratio, since this represents a system identification procedure.

Chapter 7

NUMERICAL EXAMPLE: DETECTION IN MULTIPATH CHANNELS

7.1 Introduction

This chapter has two purposes: first, to compare the performance of the optimal processor to that of suboptimal methods, and second, to verify through numerical simulations that the processing schemes proposed in Chapter 6 are correct. We have selected an example of significant practical importance; specifically, detection in a multipath propagation channel, where the signal arrives at the array in several partially or fully correlated wavefronts.

In Section 7.2, the problem is posed and the measurement model is derived in detail. Next, expressions for a meaningful receiver performance criteria are evaluated. The results will be used to compare the estimator-correlator processor with suboptimal processing techniques in Section 7.4. The performance of the generalized eigenvalue decomposition is evaluated in Section 7.5, and finally, the relationship between identification and detector performance is examined in Section 7.6.

7.2 Problem Formulation

The environment and specific array geometry which will be examined in this chapter are described in this section. A narrowband signal $s(\cdot)$ is transmitted into the medium. In a multipath propagation channel, the returning signal arrives at the array from p directions. The matrix representation of $f(\cdot)$ was derived in Section 3.4.4, and in terms of this model, the desired signal $\underline{y}(\cdot,\cdot)$ is:

$$\underline{y}(t, \omega) = \underline{fs}(t) = \sum_{k=1}^{p} b_k(\omega)\, \underline{v}_k\, s(t) \qquad (7.2\text{-}1)$$

where $\underline{v}_k$ is the steering vector associated with the kth arrival direction, and $b_k(\cdot)$ is a complex random variable representing path loss and phase shift effects.

The noise process has two components: spatially uncorrelated sensor noise, plus correlated noise from an angular spread source. The spread source is large with respect to the signal wavelength, meaning that a point scatterer representation of this component is inadequate. Its value at the ith sensor is:

$$n_i(t, \omega) = \int_{\Theta} b(\Theta, \omega)\, n(t - \tau_i(\Theta))\, d\Theta \qquad (7.2\text{-}2)$$

Assuming a narrowband $n(\cdot)$, and approximating the integral with a sum leads to the following model for the colored noise component:

$$\sum_{i=1}^{q} b_i(\omega)\, \underline{v}_i\, n(t) \qquad (7.2\text{-}3)$$

where the $\{b_i(\cdot)\}$ terms and $\{\underline{v}_i\}$ vectors are analogous to those discussed previously.

Evaluating the optimal processor requires the second order statistics of the signal process, the noise process, and the signal plus noise process as measured at the array.

The covariance matrix of $\underline{y}(\cdot,\cdot)$ at time t is:

$$\underline{R}_y = \{[\sum_{k=1}^{p} b_k \underline{v}_k s(t)] [\sum_{\ell=1}^{p} b_\ell \underline{v}_\ell s(t)]^H\}$$

$$= E\{\sum_{k=1}^{p}\sum_{\ell=1}^{p} b_k b_\ell^* \underline{v}_k \underline{v}_\ell^H |s(t)|^2\} = \sum_{k=1}^{p}\sum_{\ell=1}^{p} \underline{v}_k E\{b_k b_\ell^*\} \underline{v}_\ell^H |s(t)|^2$$

$$= \sigma_s^2 \underline{V} \underline{R}_B \underline{V}^H$$

where

$$\underline{V} = [\underline{v}_1 \ \underline{v}_2 \ \cdots \ \underline{v}_p] \tag{7.2-4}$$

is a Vandermonde matrix whose columns are the steering vectors of the arriving planar wavefronts, σ_s^2 is the signal power, and $\underline{R}_B$ is the correlation matrix of the $\{b_k(\cdot)\}$ coefficients. Of course, the wavefronts are uncorrelated if $E\{b_k \ b_\ell^*\}$ is identically zero when $k \neq \ell$.

The noise covariance matrix is

$$\underline{R}_N = \sigma_w^2 \underline{I} + \sigma_N^2 \underline{V}_N \underline{R}_{BN} \underline{V}_N^H \tag{7.2-5}$$

where σ_w^2 is the sensor noise variance, $\underline{V}_N$ is a matrix whose columns are the steering vectors $\{\underline{v}_i\}$, and $\underline{R}_{BN}$ is the covariance matrix of the path coefficients.

Finally, the signal plus noise covariance matrix $\underline{R}_1$ is the sum of $\underline{R}_N$ and $\underline{R}_y$, because we have assumed uncorrelated signal and noise components.

The optimal receiver for a snapshot of data is

$$\Lambda(\underline{r}) = \underline{r}^H(\underline{R}_N^{-1} - \underline{R}_1^{-1})\underline{r} \tag{7.2-6}$$

and the estimator-correlator form of $\Lambda(\cdot)$ is

$$\Lambda(\underline{r}) = (\underline{R}_N^{-1}\ \underline{r})^H\ \underline{R}_y\ \underline{R}_1^{-1}\ \underline{r} \tag{7.2-7}$$

This processor is seldom applied in practical array processing problems. In practice, when the arrival directions in a multipath channel are unknown, a set of closely spaced steering vectors is used to cover the sector from which the strongest return is expected. The steering vectors, multiplied by the known signal s(t), are a set of spatial filters matched against deterministic point sources immersed in an anisotropic noise field. The suboptimal log-likelihood ratio

$$\Lambda_s(\underline{r}) = \underline{r}^H\ \underline{R}_N^{-1}\ \underline{v}(\theta)\ s(t) \tag{7.2-8}$$

is computed for each steering vector, and the maximum value is used as the test statistic.

Clearly, it is easier to calculate (7.2-8) than to identify $\underline{R}_y$ and implement the optimal structure. Is optimal processing worth the added complexity? In order to answer this question, a meaningful performance criteria is needed by which the structures (7.2-7) and (7.2-8) can be compared numerically. This will be addressed in the next section.

7.3 Derivation of the Detection Index

7.3.1 Definition and Significance

We shall use the detection index as the performance criteria:

$$\Delta = \frac{E\{|\ell|^2 \mid H_1\} - E\{|\ell|^2 \mid H_0\}}{E\{|\ell|^2 \mid H_0\}} \tag{7.3-1}$$

where ℓ is the log-likelihood functional. This is a meaningful criteria for two reasons. First, the detection index can be interpreted as the output signal-to-noise ratio, which means that the processing gain of both structures can be calculated and compared. Also, it can be shown that Δ relates the detection probability to probability of false alarm:

$$P_F = P_D^{1+\Delta} \tag{7.3-2}$$

Clearly, for a given detection probability, Δ should be as large as possible to minimize the probability of false alarm. Equation (7.3-2) can be used to construct receiver operating characteristics for the estimator-correlator and the suboptimal processor. The significance of these criteria is well documented [44][47].

7.3.2 Derivation for the Suboptimal Processor

First, the suboptimal detection index will be calculated.

Equation (7.3-1) can be evaluated in terms of the channel output $\underline{s}$ and the estimated channel output $\hat{\underline{v}}$. The latter term represents the processing signal which is correlated with the filtered version of $\underline{r}(t)$.

The likelihood ratio is

$$\ell(\underline{r}) = \underline{r}^H \underline{R}_N^{-1} \hat{\underline{y}} \tag{7.3-3}$$

and

$$E\{|\ell|^2 \mid H_1\} = E\{|\underline{r}^H \underline{R}_N^{-1} \hat{\underline{y}}|^2\} = E\{(\underline{r}^H \underline{R}_N^{-1} \hat{\underline{y}})^*(\underline{r}^H \underline{R}_N^{-1} \hat{\underline{y}})\} =$$

$$E\{(\hat{\underline{y}})^H \underline{R}_N^{-1} \underline{r}\, \underline{r}^H \underline{R}_N^{-1} (\hat{\underline{y}})\} =$$

$$E\{(\hat{\underline{y}})^H \underline{R}_N^{-1} (\pounds\underline{s})(\pounds\underline{s})^H \underline{R}_N^{-1} (\hat{\underline{y}})\} + E\{(\hat{\underline{y}})^H \underline{R}_N^{-1} \underline{nn}^H \underline{R}_N^{-1} (\hat{\underline{y}})\} \tag{7.3-4}$$

The first term is:

$$E\{(\hat{\underline{y}})^H \underline{R}_N^{-1}(\pounds\underline{s})(\pounds\underline{s})^H R_N^{-1} (\hat{\underline{y}})\} = E\{|(\pounds\underline{s})^H \underline{R}_N^{-1} (\hat{\underline{y}})|^2\} \tag{7.3-5}$$

and the second is:

$$E\{(\hat{\underline{y}})^H \underline{R}_N^{-1} \underline{nn}^H (\hat{\underline{y}})\} = (\hat{\underline{y}})^H \underline{R}_N^{-1} (\hat{\underline{y}}) \tag{7.3-6}$$

Therefore,

$$E\{|\ell|^2 \mid H_1\} = E\{|(\pounds\underline{s})^H \underline{R}_N^{-1} (\hat{\underline{y}})|^2\} + (\hat{\underline{y}})^H \underline{R}_N^{-1} (\hat{\underline{y}}) \tag{7.3-7}$$

Furthermore,

$$E\{|\ell|^2 \mid H_0\} = E\{|\underline{r}^H \underline{R}_N^{-1}(\hat{\underline{y}})|^2 \mid H_0\} = E\{(\hat{\underline{y}})^H \underline{R}_N^{-1} \underline{nn}^H \underline{R}_N^{-1} (\hat{\underline{y}})\}$$

$$= (\hat{\underline{y}})^H \underline{R}_N^{-1} (\hat{\underline{y}})$$

The final result for Δ is:

$$\Delta = \frac{E\{|(\underline{\mathcal{E}s})^H \underline{R}_N^{-1}(\hat{\underline{y}})|^2\}}{(\hat{\underline{y}})^H \underline{R}_N^{-1} (\hat{\underline{y}})} \qquad (7.3\text{-}8)$$

This expression will be used to evaluate suboptimal processing performance.

The processing signal $\hat{\underline{y}}$ used in (7.3-8) is the scalar signal s(t) multiplied by a steering vector. Denote the steering vector by $\underline{v}(\theta)$. Then Equation (7.3-8) becomes

$$\Delta = \frac{E\{|(\underline{\mathcal{E}s})^H \underline{R}_N^{-1} \underline{v}(\theta)|s(t)|^2|^2\}}{\underline{v}^H(\theta) \underline{R}_N^{-1} \underline{v}(\theta) |s(t)|^2} \qquad (7.3\text{-}9)$$

Evaluating Δ in terms of the measurement model (7.2-1) and the correlation matrix $\underline{R}_B$ is useful for computation. Writing the numerator of Equation (7.3-9) in terms of the measurement model gives:

$$E\{|(\underline{\mathcal{E}s})^H \underline{R}_N^{-1} \underline{v}(\theta)\, s(t)|^2\} = E\{|[\sum_{k=1}^{p} b_k \underline{v}_k\, s(t)]^H \underline{R}_N^{-1} \underline{v}(\theta)\, s(t)|^2\}$$

Expanding this expression yields

$$E\{|[\sum_{k=1}^{p} b_k \underline{v}_k^H \underline{R}_N^{-1} \underline{v}(\theta)|^2 |s(t)|^4\} \qquad (7.3\text{-}10)$$

In order to express Equation (7.3-10) in terms of $\underline{R}_B$, define the Hermitian form

$$v_{k\theta} = \underline{v}_k^H \underline{R}_N^{-1} \underline{v}(\theta) \tag{7.3-11}$$

Then

$$E\{|\sum_{k=1}^{p} b_k \underline{v}_k^H \underline{R}_N^{-1} \underline{v}(\theta)|^2\} = E\{|\sum_{k=1}^{p} b_k v_{k\theta}|^2\} =$$

$$E\{\sum_{k=1}^{p}\sum_{\ell=1}^{p} b_k b_\ell^* v_{k\theta} v_{\ell\theta}^*\} \tag{7.3-12}$$

where, for simplicity, we have dropped the ω notation. Rearranging Equation (7.3-12) gives

$$\sum_{k=1}^{p}\sum_{k=1}^{p} E\{b_k b_\ell^*\} \underline{v}_{k\theta} v_{\ell\theta}^* = \underline{v}^H \underline{R}_B \underline{v} \sigma_s^2 \tag{7.3-13}$$

where

$$\underline{v} = [v_{1\theta} v_{2\theta} \cdots v_{p\theta}]^T \tag{7.3-14}$$

and from (7.3-11),

$$v_{i\theta} = \underline{v}_i^H \underline{R}_N^{-1} \underline{v}(\theta) \tag{7.3-14}$$

We recall that $\underline{v}_i$ is the ith steering vector. It is easy to show that the denominator of Equation (7.3-8) is

$$\underline{v}^H(\theta)\ \underline{R}_N^{-1}\ \underline{v}(\theta)\ |s(t)|^2$$

Combining these results leads to the result

$$\Delta = \frac{\underline{v}^H\ \underline{R}_B\ \underline{v}}{\underline{v}^H(\theta)\ \underline{R}_N^{-1}\ \underline{v}(\theta)}\ \sigma_s^2 \qquad (7.3\text{-}16)$$

More will be said about (7.3-16) later when specific situations are considered.

7.3.3 Derivation for the Optimal Processor

A closed form solution for the optimal structure detection index is most easily derived by expressing the log-likelihood ratio in its Hermitian form representation:

$$\ell(\underline{r}) = \underline{r}^H(\underline{R}_N^{-1} - \underline{R}_1^{-1})\underline{r} \qquad (7.3\text{-}17)$$

Since $\underline{r}$ is normally distributed under either hypothesis, evaluating $E\{|\ell|^2 \mid H_0\}$ and $E\{|\ell|^2 \mid H_1\}$ is equivalent to calculating the expectation of the product of random Hermitian forms. The solution for multivariate normal distributions is well known and available in the statistical literature [48]. Using these results leads to the following:

$$E\{|\ell|^2 \mid H_1\} = (\text{trace}\ [(\underline{R}_N^{-1} - \underline{R}_1^{-1})\ \underline{R}_1])^2$$

$$+ 2\ \text{trace}\ [(\underline{R}_N^{-1} - \underline{R}_1^{-1})\ \underline{R}_1(\underline{R}_N^{-1} - \underline{R}_1^{-1})\ \underline{R}_1] \qquad (7.3\text{-}18)$$

The arguments of the trace operator can be simplified since

$$(\underline{R}_N^{-1} - \underline{R}_1^{-1})\,\underline{R}_1 = \underline{R}_N^{-1}\,\underline{R}_1 - \underline{I} = \underline{R}_N^{-1}(\underline{R}_1 - \underline{R}_N) = \underline{R}_N^{-1}\,\underline{R}_y \quad (7.3\text{-}19)$$

Substituting (7.3-19) into (7.3-18) gives

$$E\{|\ell|^2 \mid H_1\} = [\text{trace}\,(\underline{R}_N^{-1}\,\underline{R}_y)]^2 + 2\,\text{trace}\,[(\underline{R}_N^{-1}\,\underline{R}_y)^2] \quad (7.3\text{-}20)$$

Furthermore, it is easy to show that

$$E\{|\ell|^2 \mid H_0\} = [\text{trace}\,(\underline{R}_1^{-1}\,\underline{R}_y)]^2 + 2\,\text{trace}\,[(\underline{R}_1^{-1}\,\underline{R}_y)^2] \quad (7.3\text{-}21)$$

Therefore, the detection index is evaluated by substituting these results into Equation (7.3-1).

7.4 Numerical Simulation Results

7.4.1 Experimental Description

Equations (7.3-16), (7.3-20), and (7.3-21) were evaluated numerically on a general purpose computer. The results which will be presented in this section are representative of many scenarios that have been studied.

The array consisted of eight omnidirectional sensors spaced equidistantly. The wavelength-to-spacing ratio was 0.5.

The signal and noise sources used in the simulation were both in the far field. Wavefronts which represented the desired signal arrived at the array from -5°, 0°, and 5° due to the multipath propagation channel $\mathfrak{L}(\cdot)$. The array look direction for suboptimal processing was 0°. The anisotropic spread noise source subtended a ten degree angle

and was centered at 5° with respect to array boresight. Correlation coefficients for the signal wavefronts were given by:

$$E\{b_k(\omega)b_\ell^*(\omega)\} = \exp\{-\frac{(\theta_k - \theta_\ell)^2}{150} - j\frac{2\pi(\theta_k - \theta_\ell)}{360}\} \qquad (7.4\text{-}1)$$

where θ_k is the arrival direction of the kth wavefront in degrees. This expression was used to calculate the entries in $\underline{R}_B$. It was also used to generate the entries in the noise wavefront correlation matrix $\underline{R}_{BN}$, implying that the signal and noise processes have similar statistical properties. The parameters σ_s^2, σ_N^2, and σ_W^2 were adjusted to vary the input signal-to-noise ratio.

The input signal-to-noise ratio was defined as the average signal power summed over the array divided by the average noise power over the array:

$$\text{SNR} = 10 \log (\text{trace } \underline{R}_y/\text{trace } \underline{R}_N) \qquad (7.4\text{-}2)$$

7.4.2 Optimal Versus Suboptimal Processing

Figure 7-1 illustrates processing gain as a function of input signal-to-noise ratio. These results show that optimal processing affords at least 8 dB improvement over suboptimal processing.

Receiver operating characteristics for input signal-to-noise ratios of -20 dB, -12 dB, and -9 dB are presented in Figures 7-2, 7-3, and 7-4, respectively. At very low signal-to-noise ratios (less than -20 dB), the performance of the estimator-correlator is not significantly better

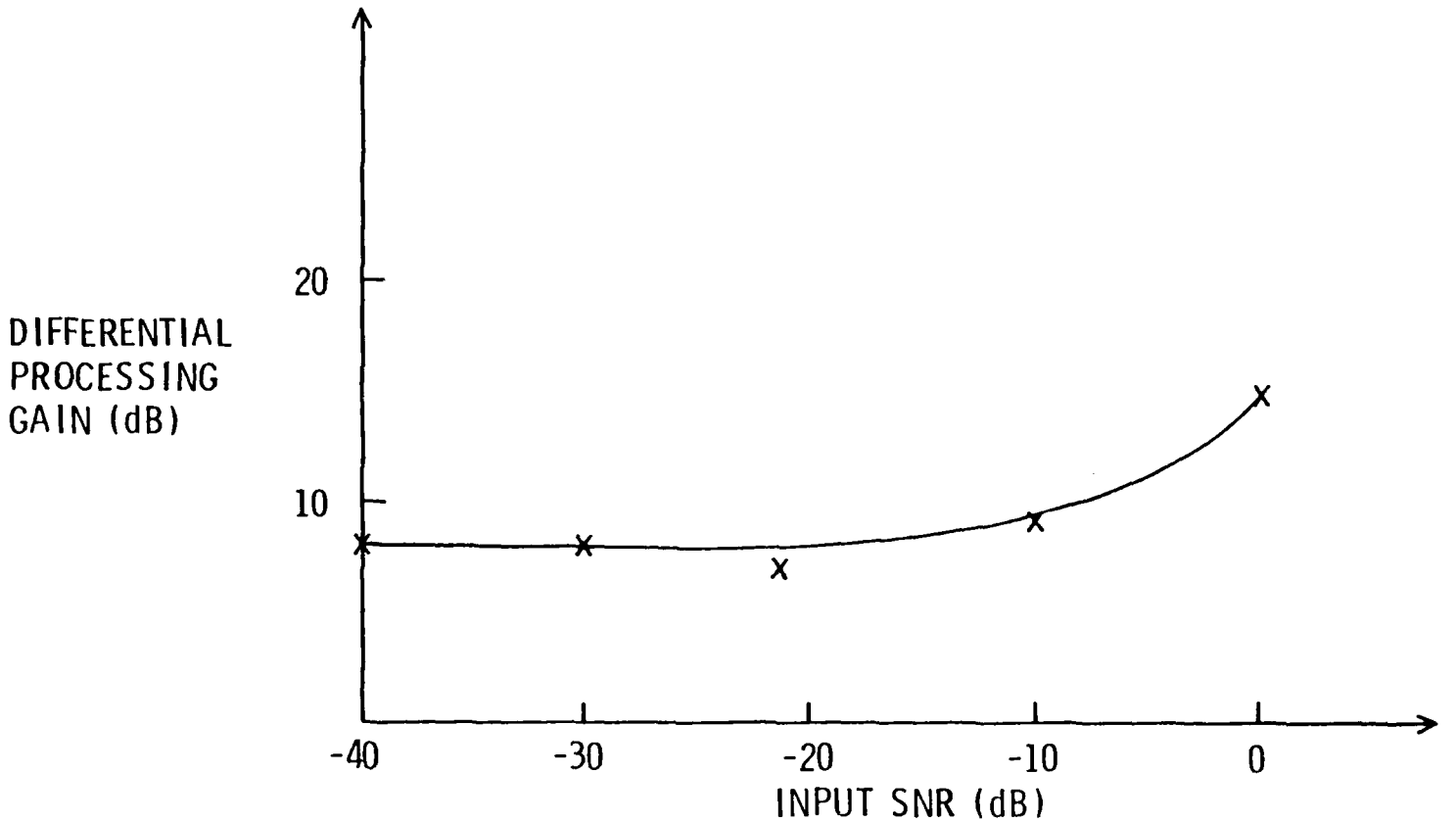


Figure 7-1. *Processing Gain Improvement*

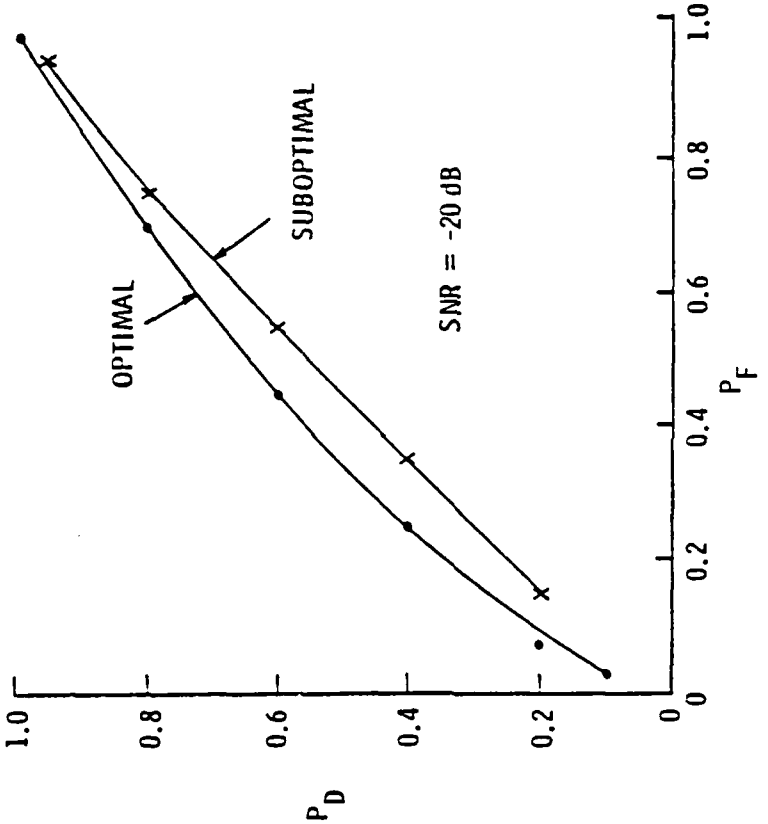


Figure 7-2. Receiver Operating Characteristic: SNR = -20 dB

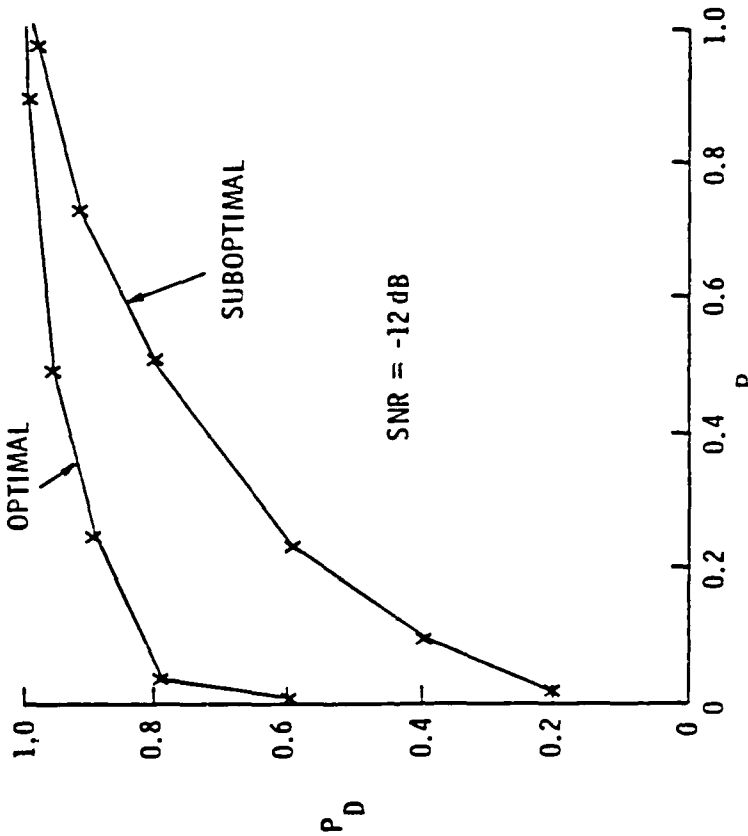


Figure 7-3. Receiver Operating Characteristic: SNR = -12 dB

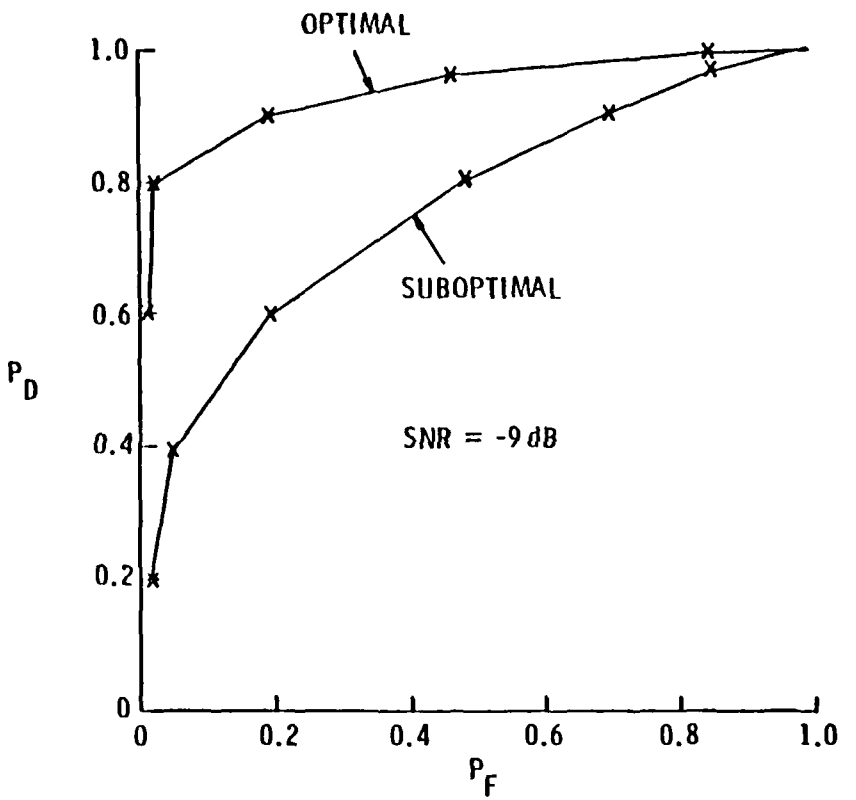


Figure 7-4. Receiver Operating Characteristic: SNR = -9 dB

than the suboptimal processor, because in both cases, Δ is small, and the sum $1 + \Delta$ is approximately one. Therefore, the eight dB processing gain is not translated into an improvement in detection probability for a given false alarm rate. However, once the SNR is greater than -20 dB, improvement with respect to this criteria can be clearly seen.

7.5 Calculating the Estimator Kernel

7.5.1 Introduction

The second half of this study actually has two purposes: first, to demonstrate that the generalized singular value decomposition can be used to construct $\underline{G}(\cdot,\cdot)$ directly from data, and second, to determine what signal-to-noise ratio is needed to estimate $\underline{R}_y(\cdot,\cdot)$ (or to form data matrix $\underline{B}$). This is an important parameter. Of course, the probing signal is under our control, therefore in principle, the input SNR could be raised to any desired magnitude. However, in practical situations this is clearly not possible; therefore, the ability to estimate $\underline{R}_y(\cdot,\cdot)$ and $\mathfrak{L}(\cdot)$ at a moderate SNR is highly desirable. If an input SNR of 80 dB is needed for identification, one can safely conclude that this approach to adaptive implementation is impractical!

7.5.2 Experimental Description

The multipath channel model, noise models, and array model are identical to those presented in Sections 7.2 and 7.4.1. Both $\underline{R}_y$ and $\underline{R}_N$ were normalized so that each had unit trace, meaning that the nominal input SNR was 0 dB.

In order to generate simulated array measurements, a standard IMSL routine was used to first generate sets of white Gaussian noise. Next, the white noise was formed into random vectors, and filtered by the square roots of $\underline{R}_y$ and $\underline{R}_N$ in order to construct random vectors. After filtering, the random vectors represented array data given signal alone and noise alone, respectively. Noisy measurements of the channel output were constructed by scaling the signal alone vectors and adding them to the noise alone vectors.

7.5.3 Results

Several generalized singular value decomposition algorithms were made available to the author courtesy of Charles Van Loan of Cornell University. After modifying several of them for complex-valued matrices, they were tested in two ways. First, we attempted to calculate $\underline{C}$ from the square roots of the actual covariance matrices $\underline{R}_1$ and $\underline{R}_y$. To do this, their eigenvalue decompositions were computed, and matrices $\underline{A}$ and $\underline{B}$ were defined as follows:

$$\underline{A} = \underline{R}_1^{1/2} = \underline{\Lambda}_1^{1/2} \underline{U}_1 \tag{7.5-1}$$

$$\underline{B} = \underline{R}_y^{1/2} = \underline{\Lambda}_y^{1/2} \underline{U}_y \tag{7.5-2}$$

It is clear from Equations (7.5-1) and (7.5-2) that

$$\underline{A}^H \underline{A} = \underline{R}_1 \tag{7.5-3}$$

$$\underline{B}^H \underline{B} = \underline{R}_y \tag{7.5-4}$$

By following the procedure in Section 6.3.3, it should be possible to calculate $\underline{G}$ from (7.5-1) and (7.5-2).

We found that the procedure discussed above worked perfectly. The result was compared with

$$\underline{G} = \underline{R}_y \, \underline{R}_1^{-1}$$

and the answers matched exactly. We concluded that the algorithms worked correctly, and that it should be possible to compute $\underline{G}$ from data.

Next, we attempted to calculate $\underline{G}(\cdot,\cdot)$ directly from array data matrices. Data matrix $\underline{A}$ was formed from 256 array measurements given a priori knowledge that H_1 was true. The data matrix $\underline{B}$ was formed by increasing the probing signal power, taking 256 array measurements, and then scaling the matrix by a factor of $1/\sigma_s$. This step normalized the trace of $\hat{\underline{R}}_y$ to the proper value, which in this example was unity. The dimensions of $\underline{A}$ and $\underline{B}$ were 256 by 8. These matrices were input into the generalized singular value decomposition algorithm, and the processing steps described in Section 6.3.3 were followed in an attempt to compute $\underline{G}(\cdot,\cdot)$.

Unfortunately, these attempts were unsuccessful. The data matrix $\underline{B}$ was constructed at signal-to-noise ratios ranging from eight to 25 dB, and in all cases, the algorithm was unstable. We do not know why the algorithms did not work during these trials; however, since the algorithms are very new, it is likely that the computer programs have not been perfected.

Another approach to the problem was tried. Since the algorithms worked for the square roots of $\underline{R}_1$ and $\underline{R}_y$, we computed the covariance matrix estimates $\hat{\underline{R}}_1$ and $\hat{\underline{R}}_y$ by

$$\hat{\underline{R}}_1 = \underline{A}^H \underline{A}$$

$$\hat{\underline{R}}_y = \underline{B}^H \underline{B}$$

Of course, this procedure defeats the reason why the CS decomposition was proposed in the first place, but we wanted to demonstrate in principle that $\underline{G}(\cdot,\cdot)$ could be constructed from data! Next, eigenvalue decompositions of $\hat{\underline{R}}_1$ and $\hat{\underline{R}}_y$ were computed, and their square roots formed. The processing procedure in Section 6.3.3 was carried out using the square roots of the estimated covariance matrices in another attempt to obtain $\underline{G}(\cdot,\cdot)$ from array data.

This approach was successful, and furthermore, accurate estimates of $\underline{R}_y$ could be made at moderate signal-to-noise ratios. An input SNR of 15 dB was sufficient for excellent identification. This judgement was empirical, because we compared the actual $\underline{R}_y$ to the estimated matrix element by element. As the input SNR was increased, we found little improvement after an SNR of 15 dB was attained. We concluded that the identification scheme is feasible, and can be carried out at a moderate SNR. Although it is not possible to generalize these results with absolute certainty, it seems reasonable to conclude that they carry over to other channel models.

7.6 Relating Identification to Processor Performance

The results from Section 7.5 demonstrated that $\underline{R}_y(\cdot,\cdot)$ could be identified accurately at moderate signal-to-noise ratios. However, they did not made a connection between identification and optimal processor performance. Therefore, further experiments were conducted in order to determine how well $\underline{R}_y(\cdot,\cdot)$ must be identified to obtain processing gain improvement as compared to suboptimal methods.

The channel, noise, and array models were identical to those presented in Sections 7.2, 7.4.1, and 7.5.2. The optimal detection index Δ (Equation (7.3-1)) was evaluated for $\underline{R}_y(\cdot,\cdot)$ identified at low and high signal-to-noise ratios. Closed-form expressions for $E\{|\ell|^2|H_1\}$ and $E\{|\ell|^2|H_0\}$ given misidentified $\underline{R}_y(\cdot,\cdot)$ can be obtained; however, they are difficult to evaluate analytically. Therefore, Monte Carlo methods were used to evaluate Δ.

The results of our experiments are illustrated in Figure 7-5. They are interesting and intuitively pleasing. As the signal-to-noise ratio increases, the processing gain approaches the theoretical predicted maximum. Little processing improvement is achieved above an identification SNR of 15 dB, which makes the connection between receiver performance and the empirical observation made in the previous section. However, the most significant conclusion of this experiment is that even a poor identification results in improved detection. Perfect identification at high signal-to-noise ratios is not needed.

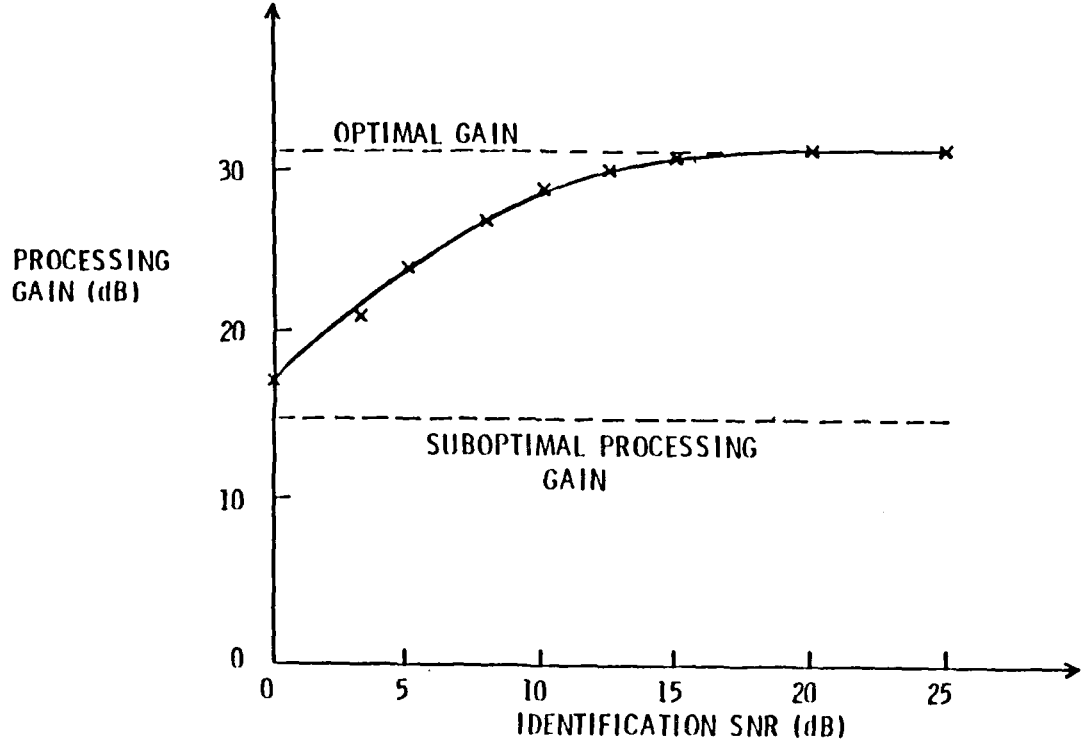


Figure 7-5. Relating Identification to Processing Performance

7.7 Conclusions

Matrix representations can be used to model non-trivial propagation and scattering channels, and are a useful tool for computer simulation experiments. In this chapter, they were used successfully to evaluate the estimator-correlator processor, compare its performance against a suboptimal receiver, and to generate simulated array measurements for testing new computational algorithms.

The relationship between identification and receiver performance was examined. We found that perfect identification was not required to improve processing gain. Even a poor identification conducted at a low SNR resulted in some improvement.

Also, the performance of optimal versus suboptimal processing was evaluated numerically. It was found that optimal processing gives at least eight dB improvement over suboptimal techniques. For input signal-to-noise ratios greater than -20 dB, this translated into an improved receiver operating characteristic.

The performance of new generalized singular value decomposition algorithms was evaluated. We had some difficulties with them, yet were able to show that $\underline{C}(\cdot,\cdot)$ could be computed in principle. There is nothing incorrect with the proposed processing scheme. The programs which were tested are new, and require further testing and debugging.

Chapter 8

CONCLUDING REMARKS

8.1 Conclusions

The estimator-correlator processor establishes fundamental connections among detection theory, estimation theory, system modeling, and system identification theory. Variations of this canonical structure solve the detection problem for the generalized exponential class of signal and noise distributions. This result establishes a basic connection between detection and estimation theory.

An operator theoretic approach to the channel representation problem allows the detection and modeling problems to be solved for a very wide class of transmission media, especially spread channels, which are particularly difficult to handle. Moreover, this systematic approach made the full power of Hilbert space theory and functional analysis available for use in subsequent derivations.

Matrix representations of bounded, linear operators are useful for modeling a wide range of deterministic and stochastic transformations one might encounter in practical array processing problems. They are easily incorporated into the estimator-correlator structure. Identification of $\mathfrak{L}(\cdot)$ as represented by matrices is needed in order to calculate the conditional mean of the channel output $\mathfrak{L}\underline{s}(t)$. Measuring the matrix element cross-correlations represents a systematic approach to stochastic Green's function identification.

The channel identification problem can be simplified through simultaneous diagonalization of the input and output covariance kernels.

Simultaneous diagonalization can be accomplished either through signal design or generalized singular value decomposition. The result establishes an interesting connection among detection, estimation, system modeling, and identification, and in addition, provides new insight into classical system identification issues. The Karhunen-Loève expansion furnishes a fundamental structure for stochastic system modeling and identification.

Solving the space-time processor equations through orthogonal decompositions represents the most important accomplishment of this dissertation. Karhunen-Loève representations are the key to both theoretical analysis and adaptive implementation. Important connections between this expansion and other decompositions, including the singular value decomposition, CS decompositions, generalized eigenvalue factorizations, QR factorizations, and generalized Fourier series have been made. Combinations of these canonical decompositions and orthonormal representations provide the key to implementing the processor with numerically robust algorithms.

The array processing algorithms proposed in this dissertation are more than academic ideas that can not work in practice. They were thoroughly tested and they work well. The computational burden is worth the effort, because optimal processing is significantly better than simpler suboptimal techniques. Identification can be carried out at moderate signal-to-noise ratios. Moreover, perfect channel identification is unnecessary. Even a poor identification improves receiver performance.

8.2 Recommendations for Future Research

The numerical problems involved with implementing the estimator-correlator must be examined. It is not known how finite-precision word lengths, matrix ill-conditioning, and channel misidentification work together to affect the performance of the estimator-correlator. Understanding these effects is crucial in order to build the processor in hardware.

Several theoretical issues warrant further research. For example, some recent work suggests that the structure of $\underline{L}(\cdot)$ gives insight into "how far" a stochastic system deviates from stationarity. This idea needs to be developed, since it has the potential to give new insights into stochastic system characterization.

In this dissertation, the theoretical foundations of optimal space-time array processing have been examined at length. We demonstrated how the processing equations can be solved and implemented with robust numerical algorithms. It is clear from Chapter 6 that implementing the processor is very demanding computationally. However, the data flow and computations are regular, repetitive, and well suited for parallel computations by distributed processors. Algorithms for performing singular value decompositions and CS decompositions that are amenable to parallel processing or systolic array implementation need to be developed. Solving these problems will require new basic computational cells and new methods to assess their computational complexity.

Developing new basic computational cells will require a deeper understanding of the fundamental structure of computational algorithms.

For example, a basic algorithm used in many signal processing operations is the generalized coordinate rotation [49]. In many applications, it can be regarded as more fundamental than even the traditional floating point operation (complex multiply-and-accumulate). Surprisingly enough, the generalized coordinate rotation algorithm has a fundamental connection with Lie group theory, an abstract mathematical discipline [50]. More work in this area is needed to establish deeper connections between Lie group theory and basic signal processing algorithms.

The connection between the work presented in this dissertation and other basic research areas in adaptive signal processing is presented in Figure 8-1. The overall effort calls on disciplines such as Lie group theory, graph theory, and information-theoretic analysis of computational complexity, in addition to stochastic operator theory, numerical analysis, and integral equation theory. All of these tools contribute to the understanding needed for efficient optimal array processor design.

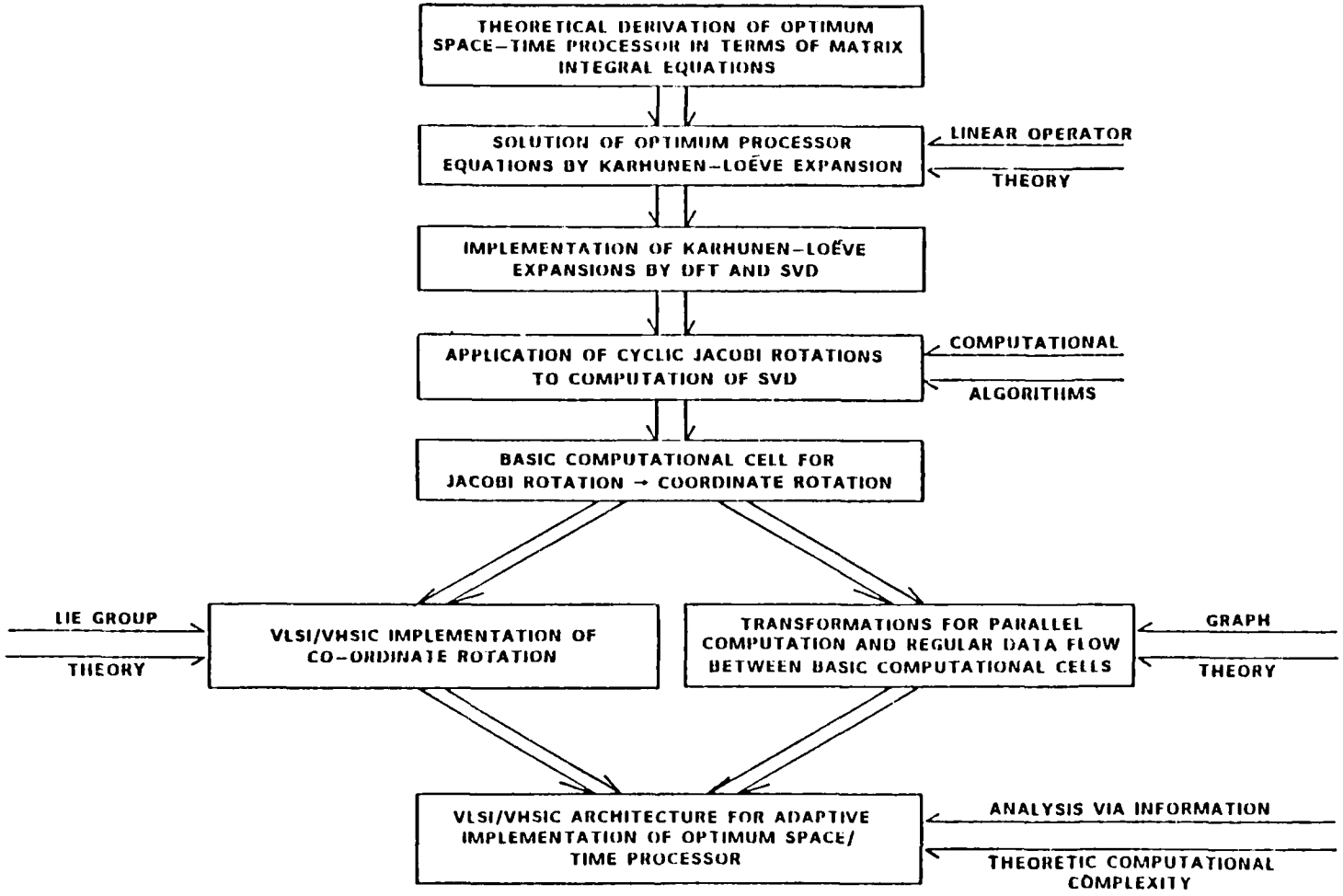


Figure 8-1. Interaction of Research Topics in Adaptive Signal Processing

APPENDIX

CONTOUR INTEGRAL EVALUATION

Calculating the matrix representations for the operators appearing in Sections 3.2.3, 3.2.5, and 3.2.6 requires evaluating three definite integrals.

The representation for the time delay operator is:

$$a_{mn} = \langle A_\tau \phi_m, \phi_n \rangle = \frac{1}{T} \int_{-\infty}^{+\infty} \frac{\sin \sigma(t - \tau - mT)}{\sigma(t - \tau - mT)} \frac{\sin \sigma(t - nT)}{\sigma(t - nT)} \, dt \quad \text{(A-1)}$$

To evaluate Equation (A-1), begin by simplifying the product in the numerator using the trigonometric identity

$$\sin \alpha \sin \beta = \frac{1}{2} \cos(\alpha - \beta) - \frac{1}{2} \cos(\alpha + \beta)$$

By setting $\alpha = \sigma\tau + m\sigma T$ and $b = n\sigma T$ then

$$\sin \sigma(t - \tau - mT) \sin \sigma(t - nT) = \sin(\sigma T - \alpha) \sin(\sigma t - b) =$$

$$\frac{1}{2} \cos(b - a) - \frac{1}{2} \cos(2\sigma t - (a + b))$$

and (A-1) becomes

$$a_{mn} = \frac{1}{2T} \int_{-\infty}^{+\infty} \frac{\cos(b - a)}{(\sigma t - a)(\sigma t - b)} \, dt - \frac{1}{2T} \int_{-\infty}^{+\infty} \frac{\cos(2\sigma t - (a + b))}{(\sigma t - a)(\sigma t - b)} \, dt \quad \text{(A-2)}$$

Evaluate the first integral in (A-2):

$$\frac{1}{2T}\int_{-\infty}^{+\infty}\frac{\cos(b-a)}{(\sigma t-a)(\sigma t-b)}\,dt=\frac{\cos(b-a)}{2T\sigma^2}\int_{-\infty}^{+\infty}\frac{dt}{(t-a/\sigma)(t-b/\sigma)} \quad (A-3)$$

Convert (A-3) into a contour integral on the z-plane:

$$\frac{\cos(b-a)}{2T\sigma^2}\int_C\frac{dz}{(z-a/\sigma)(z-b/\sigma)} \quad (A-4)$$

where C is shown in Figure A-1. Reference [51] shows that

$$\int_C\frac{dz}{(z-a/\sigma)(z-b/\sigma)}=\int_C f(z)\,dz=\pi i\sum \text{Res } f(z)$$

with Res f(z) meaning the residue of f(z). The residue at $x = a/\sigma$ is

$$\frac{1}{a/\sigma-b/\sigma}=\frac{\sigma}{(a-b)}$$

The residue at $x = b/\sigma$ is

$$\frac{1}{b/\sigma-a/\sigma}=\frac{\sigma}{(b-a)}$$

Therefore,

$$\int_c\frac{dz}{(z-a/\sigma)(z-b/\sigma)}=0$$

The second integral shall be calculated in a similar manner:

$$\int_{-\infty}^{+\infty}\frac{\cos(2\sigma t-(a+b))}{(\sigma t-a)(\sigma t-b)}\,dt=\text{Re}\left\{\int_{-\infty}^{+\infty}\frac{\exp(i(2\sigma t-(a+b))}{(\sigma t-a)(\sigma t-b)}\,dt\right\} \quad (A-5)$$

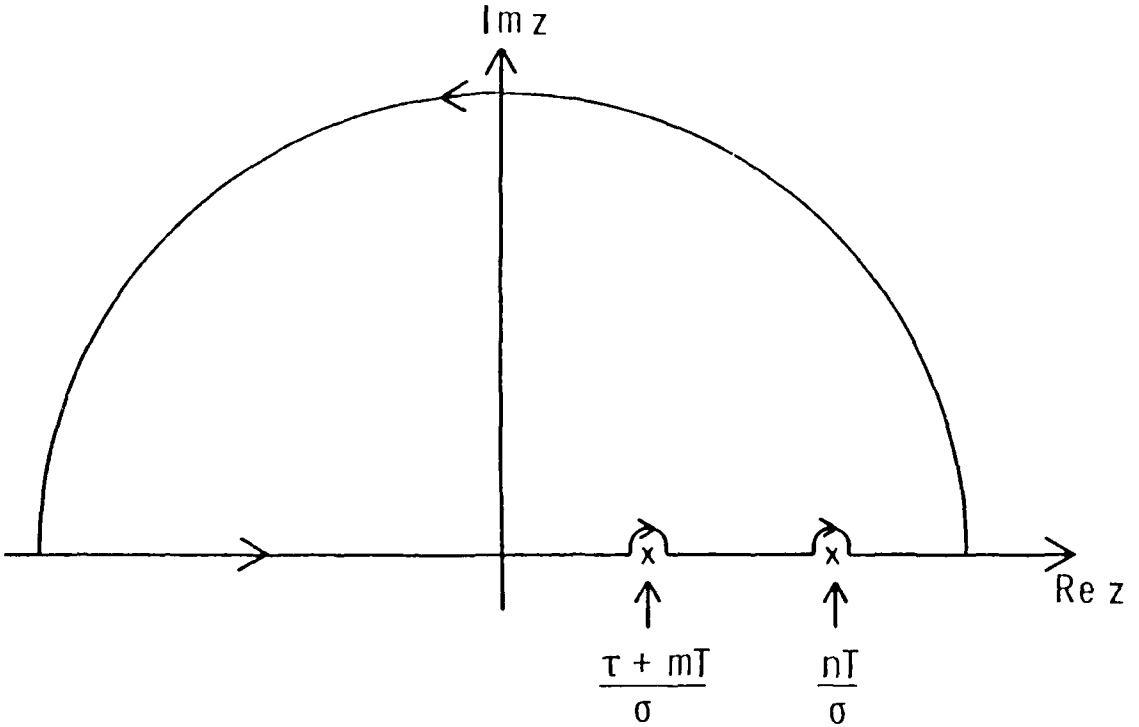


Figure A-1. Integration Contour

Convert (A-5) into a contour integral:

$$\int_C \frac{\exp(i(2\sigma z - (a + b))}{(\sigma z - a)(\sigma z - b)}\, dz = \pi i \sum \text{Res } f(z)$$

where

$$f(z) = \frac{\exp(i(2\sigma z - (a + b))}{(\sigma z - a)(\sigma z - b)} \qquad \text{(A-6)}$$

The residue of f at a/σ is

$$\frac{\sigma}{a - b}\, e^{i(a-b)}$$

and at $z = b/\sigma$ is

$$\frac{\sigma}{b - a}\, e^{i(b-a)}$$

Then, summing the residues and multiplying the result by πi gives

$$\frac{2\pi\sigma}{b - a}\, \frac{e^{i(a-b)} - e^{-i(a-b)}}{2i} = \frac{2\pi\sigma}{b - a} \sin(a - b) \qquad \text{(A-7)}$$

Almost finished! Returning to Equation (A-1) gives:

$$a_{mn} = -\frac{1}{2T} \int_{-\infty}^{+\infty} \frac{\cos(2\sigma t - (a + b))}{(\sigma t - a)(\sigma t - b)}\, dt = -\frac{1}{2T}\, \frac{2\pi\sigma}{(b - a)} \sin(a - b)$$

$$= \frac{\sin(a - b)}{(a - b)}$$

Since

$$a - b = \sigma\tau - m\sigma T - n\sigma T = \sigma(\tau - (n - m)T)$$

the coefficients are

$$a_{nm} = \frac{\sin\sigma(\tau - (n - m)T)}{\sigma(\tau - (n - m)T)} \qquad \text{(A-8)}$$

which is the answer.

The matrix representation of the stretching/compression operator is given by:

$$\frac{1}{T} \int_{-\infty}^{+\infty} \frac{\sin\, \sigma(\alpha t - mT)}{\sigma(\alpha t - mT} \, \frac{\sin\, \sigma(t - nT)}{\sigma(t - nT)} \, dt \qquad \text{(A-9)}$$

for $\alpha > 0$.

The calculation is straightforward and similar to the first example. The answer is:

$$a_{nm} = \frac{\sin(\sigma T(\alpha n - m)/\alpha)}{(\sigma T(\alpha n - m)/\alpha)} \qquad \text{(A-10)}$$

To check the answer, if $\alpha = 1$, then

$$a_{nm} = \langle \phi_n, \phi_m \rangle = \delta_{mn}$$

which is the identity operator.

Finally, the representation of the combined delay and stretching operator is obtained by evaluating

$$\frac{1}{T}\int_{-\infty}^{\infty}\frac{\sin\sigma(\alpha t-\tau-mT)}{\sigma(\alpha t-\tau-mT)}\,\frac{\sin\sigma(t-nT)}{\sigma(t-nT)}\,dt$$

using the method of residues, with the result:

$$a_{nm}=\frac{\sin\sigma(\alpha\tau-(an-m)T/\alpha)}{\sigma(\alpha\tau-(an-m)T/a)}$$

BIBLIOGRAPHY

[1] H. L. Van Trees, Detection, Estimation, and Modulation Theory (Part Three). New York: John Wiley and Sons, 1971.

[2] H. L. Van Trees, "A unified theory for optimum array processing," Arthur D. Little, Inc. Report No. 4160866, August 1966.

[3] D. Middleton, Introduction to Statistical Communication Theory. New York: McGraw-Hill, 1960.

[4] G. R. L. Sohie, "Applications of Hilbert Space Theory to Optimal and Adaptive Space-Time Processing," Ph.D. Dissertation, The Pennsylvania State University, 1983.

[5] L. H. Sibul, "Optimum Detection and Estimation in Stochastic Transmission Media," in Applied Stochastic Processes, G. Adomian, editor. New York: Academic Press, 1980.

[6] R. Price, "Optimum detection of random signals in noise with applications to scatter-multipath communications, I," IRE Transactions on Information Theory, vol. IT-2, pp. 125-135, December 1956.

[7] T. Kailath, "A note on least-squares estimates from likelihood ratios," Information and Control, vol. 13, pp. 534-540, 1968.

[8] R. Esposito, "A class of estimators for optimum adaptive detection," Information and Control, vol. 10, pp. 137-148, 1967.

[9] T. Kailath, "A general likelihood-ratio formula for random signals in Gaussian noise," IEEE Transactions on Information Theory, vol. IT-15, pp. 350-361, May 1969.

[10] S. C. Schwartz, "The estimator-correlator for discrete-time problems," IEEE Transactions on Information Theory, vol. IT-23, pp. 93-100, January 1977.

[11] R. N. McDonough, "A canonical form of the likelihood detector for Gaussian random vectors," Journal of the Acoustical Society of America, vol. 49, pp. 402-406, 1971.

[12] F. Bryn, "Optimum signal processing of three-dimensional arrays operating on Gaussian signals and noise," Journal of the Acoustical Society of America, vol. 34, pp. 289-297, March 1962.

[13] D. J. Edelblute, J. M. Fisk, and G. L. Kinneson, "Criteria for optimum signal detection theory for arrays," Journal of the Acoustical Society of America, vol. 41, pp. 199-206, January 1967.

[14] H. Cox, "Interrelated problems in estimation and detection, I and II," NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics, Enchede, The Netherlands, pp. 1-64, August 1968.

[15] R. A. Monzingo and T. W. Miller, Introduction to Adaptive Arrays. New York: John Wiley and Sons, 1980.

[16] S. Haykin, editor, Array Processing Applications to Radar. Stroudsburg, PA: Dowden, Hutchinson, and Ross, Inc., 1980.

[17] J. E. Hudson, Adaptive Array Principles. New York: Institution of Electrical Engineers, 1981.

[18] B. Widrow and S. D. Stearns, Adaptive Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[19] L. H. Sibul, editor, Adaptive Signal Processing. New York: IEEE Press, 1987.

[20] F. Riesz and B. Sz.-Nagy, Functional Analysis. Translated by L. F. Boron. New York: F. Ungar Publishing, 1955.

[21] A. W. Naylor and G. R. Sell, Linear Operator Theory in Engineering and Science. New York: Springer-Verlag, 1982.

[22] I. Gohberg and S. Goldberg, Basic Operator Theory. Boston: Birkhäuser, 1981.

[23] A. Papoulis, Probability, Random Variables, and Stochastic Processes (Second Edition). New York: McGraw-Hill, 1984.

[24] P. S. Maybeck, Stochastic Models, Estimation, and Control (Volume One). Orlando: Academic Press, 1979.

[25] W. B. Davenport and W. L. Root, An Introduction to the Theory of Random Signals and Noise. New York: McGraw-Hill, 1958.

[26] T. Kailath, "A view of three decades of linear filtering theory," IEEE Transactions on Information Theory, vol. IT-20, pp. 146-180, March 1974.

[27] S. C. Liu and L. W. Nolte, "Performance evaluation of array processors for detecting Gaussian acoustic signals," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-28, pp. 328-333, June 1980.

[28] W. S. Hodgkiss and L. W. Nolte, "Optimal array processor performance trade-offs under directional uncertainty," IEEE Transactions on Aerospace and Electronic Systems, vol. AES-12, pp. 605-615, September 1976.

[29] L. A. Zadeh and C. A. Desoer, Linear System Theory: The State Space Approach. New York: McGraw-Hill, 1963.

[30] G. Adomian, Stochastic Systems. New York: Academic Press, 1983.

[31] K. Sobczyk, Stochastic Wave Propagation. New York: Elsevier Science Publishing Co., 1985.

[32] L. H. Sibul and J. A. Tague, "Matrix representations of propagation and scattering operators," Proceedings of the 1986 Conference on Information Sciences and Systems, Princeton Unviersity, Princeton, NJ, March 1986.

[33] A. Paulraj and T. Kailath, "Beamforming in the presence of multipath," Proceedings of the 1985 IEEE Conference on Acoustics, Speech, and Signal Processing, Tampa, FL, April 1985.

[34] L. J. Ziomek, Underwater Acoustics - A Linear Systems Theory Approach. Orlando, FL: Academic Press, 1985.

[35] T. Kailath, S. Y. Kung, and M. Morf, "Displacement ranks of matrices and linear equations," Journal of Mathematical Analysis and Applications, vol. 68, pp. 395-407, 1979.

[36] T. Kailath, S. Y. Kung, and M. Morf, "Displacement ranks of a matrix," Bulletin of the American Mathematical Society, vol. 1, pp. 767-773, 1979.

[37] P. Eykhoff, System Identification. New York: John Wiley and Sons, 1974.

[38] K. J. Åström and P. Eykhoff, "System identification - a survey," Automatica, vol. 7, pp. 123-162, 1971.

[39] G. H. Golub and C. F. Van Loan, Matrix Computations. Baltimore: The Johns Hopkins University Press, 1983.

[40] N. J. Nilsson, Learning Machines. New York: McGraw-Hill, 1965.

[41] C. F. Van Loan and J. Speiser, "Computation of the CS decomposition, with application to signal processing," SPIE Volume 696 - Advanced Algorithms and Architectures for Signal Processing, pp. 71-77, 1986.

[42] G. W. Stewart, Introduction to Matrix Computations. New York: Academic Press, 1973.

[43] R. A. Horn and C. R. Johnson, Matrix Analysis. Cambridge: Cambridge University Press, 1985.

[44] A. D. Whalen, Detection of Signals in Noise. New York: Academic Press, 1971.

[45] C. F. Van Loan, "Generalizing the singular value decomposition," SIAM Journal on Numerical Analysis, vol. 13, pp. 76-83, March 1976.

[46] R. Kumaresan and D. W. Tufts, "Singular value decomposition and spectral analysis," Proceedings of the First ASSP Workshop on Spectral Estimation, pp. 6.4.1-6.4.12, August 1981.

[47] H. L. Van Trees, Detection, Estimation, and Modulation Theory (Part One). New York: John Wiley and Sons, 1968.

[48] F. A. Graybill, Matrices With Applications in Statistics. Belmont, CA: Wadsworth International Group, 1983.

[49] J. E. Volder, "The CORDIC trigonometric computing technique," IRE Transactions on Electronic Computers, vol. EC-8, pp. 330-334, September 1959.

[50] D. T. L. Lee and M. Morf, "Recursive square root estimation algorithms," Proceedings of the 1980 IEEE Conference on Acoustics, Speech, and Signal Processing, Denver, CO, April 1980.

[51] L. A. Rubenfeld, A First Course in Applied Complex Variables. New York: John Wiley and Sons, 1985.

VITA

John A. Tague was born on September 27, 1953, in Waukegan, Illinois. He earned the Bachelor of Science degree in Electrical Engineering from the University of Illinois in December 1975, and served in the U.S. Navy for five years. Mr. Tague earned the Ph.D. degree in Electrical Engineering from The Pennsylvania State University. While attending Penn State, he served as a Teaching Assistant in the Department of Electrical Engineering, and as a Graduate Assistant at the Applied Research Laboratory. Mr. Tague's research interests are adaptive array processing, numerical signal processing, and detection theory.

Mr. Tague is a member of the Institute of Electrical and Electronic Engineers, and the Society for Industrial and Applied Mathematics.

END

DATE

FILMED

MARCH

1988

DTIC